

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2001, 09:54:43 ; Search time 15.24 Seconds
(without alignments)
317.770 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 747

Sequence: 1 PPOPSRPPAPPAPPSPALPR.....VNSTWRTVRLSATACGCLG 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 94743 seqs, 34840360 residues

Total number of hits satisfying chosen parameters: 94743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 244 | 32.7 | 156 | 1 | PSPN_HUMAN |
| 2 | 241 | 32.3 | 197 | 1 | NRTN_HUMAN |
| 3 | 232.5 | 31.1 | 156 | 1 | PSPN_MOUSE |
| 4 | 231 | 30.9 | 195 | 1 | NRTN_MOUSE |
| 5 | 221.5 | 29.7 | 156 | 1 | PSPN_RAT |
| 6 | 187.5 | 25.1 | 211 | 1 | GNF_MOUSE |
| 7 | 182.5 | 24.4 | 211 | 1 | GNF_RAT |
| 8 | 175.5 | 23.5 | 211 | 1 | GNF_HUMAN |
| 9 | 116.5 | 15.6 | 436 | 1 | GDF6_BOVIN |
| 10 | 112.5 | 15.1 | 575 | 1 | MIS_BOVIN |
| 11 | 110.5 | 14.8 | 560 | 1 | MIS_HUMAN |
| 12 | 107 | 14.3 | 553 | 1 | MIS_RAT |
| 13 | 105.5 | 14.1 | 303 | 1 | GDF6_RAT |
| 14 | 103.5 | 13.9 | 303 | 1 | GDF6_MOUSE |
| 15 | 102 | 13.7 | 308 | 1 | GDF6_HUMAN |
| 16 | 102 | 13.7 | 1733 | 1 | VNUA_PRRKA |
| 17 | 101.5 | 13.6 | 555 | 1 | MIS_MOUSE |
| 18 | 99 | 13.3 | 575 | 1 | MIS_PIG |
| 19 | 97.5 | 13.1 | 485 | 1 | SSGP_VOLCA |
| 20 | 96.5 | 12.9 | 775 | 1 | ICP0_HSV1 |
| 21 | 96 | 12.9 | 350 | 1 | DAF7_CAEEL |
| 22 | 94.5 | 12.7 | 297 | 1 | HXIM_CHICK |
| 23 | 94.5 | 12.7 | 992 | 1 | AXN1_MOUSE |
| 24 | 94 | 12.6 | 376 | 1 | VASP_MOUSE |
| 25 | 94 | 12.6 | 992 | 1 | POLS_RUBVM |
| 26 | 92.5 | 12.4 | 473 | 1 | ULL3_HCMVA |
| 27 | 92 | 12.3 | 151 | 1 | GDF7_MOUSE |
| 28 | 92 | 12.3 | 309 | 1 | HXA4_CHICK |
| 29 | 92 | 12.3 | 361 | 1 | IHA_TRIVU |
| 30 | 92 | 12.3 | 629 | 1 | YS50_MYCTU |
| 31 | 92 | 12.3 | 641 | 1 | EBN1_EBV |
| 32 | 91.5 | 12.2 | 551 | 1 | EBV4_HUMAN |
| 33 | 91 | 12.2 | 2205 | 1 | POLN_RUBVT |

P13889 rubella vir

Query Match

32.7% ; Score 244 ; DB 1 ; Length 156;

ALIGNMENTS

RESULT 1
PSPN_HUMAN

ID PSPN_HUMAN STANDARD; PRT; 156 AA.

AC 060542;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE PERSEPHIN PRECURSOR (PSP).

GN PSPN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98150950; PubMed=9491986;

RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Balch R.H., Leitner M.L.,

RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbaue P.T.,

RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,

RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,

RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,

RA Phillips H.S., Johnson E.M.;

RT "Persephin, a novel neurotrophic factor related to GDNF and

RT neuritin";

RL Neuron 20:245-253(1998).

CC -!- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC

CC DOPAMINERGIC AND MOTOR NEURONS.

CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF040962; AAC39640.1 ;

DR HSP; Q07731; IAGQ.

DR MIM; 602921 ;

DR InterPro; IPR001839; TGF-beta.

DR SMART; SM00204; TGF; 1.

DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.

KW Growth factor; Signal.

FT SIGNAL 1 21

FT CHAIN 22 156

FT DISULFID 66 124

FT DISULFID 93 152

FT DISULFID 97 154

FT DISULFID 123 123

SQ SEQUENCE 156 AA; 16600 MW; 6547751653A7044A CRC64;

POTENTIAL.

PERSEPHIN.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

INTERCHAIN (BY SIMILARITY).

Best Local Similarity 44.8%; Pred. No. 7.8e-12;
Matches 56; Conservative 17; Mismatches 36; Indels 16; Gaps 3;

QY 23 RAARAGG-----PGRARAAGCRLRSOLVPRALGLGHRSDLVRFRCGSGC-R 74
DB 40 QVAKAGGTWLTGTHRLPLRRLRSLGQWLSUTLSVAELGLGYASEEKVIFRYCAGSGCP 99
QY 75 RARSPHDLSLALGAGALRPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATA 134
DB 100 GARTQGLALARLOGG-----RAHGGPCCRTRYTDVAFLDDRHWQLPQLSAAA 151
QY 135 CGCLG 139
DB 152 CGCGG 156

RESULT 2
NRTN_HUMAN STANDARD: PRT; 197 AA.
AC Q9748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEURTURIN PRECURSOR.
GN NRTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=97100947; PubMed=8945474;
RX Kotsbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,
RA Creedon P.J., Johnson E.M. Jr., Milbrandt J.;
RT "Neurturin, a relative of glial cell-line-derived neurotrophic
RT factor";
RL Nature 384:467-470(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Blum H., Bauersachs S., Meves H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP VARIANT HSCR SER-96.
RC TISSUE=Peripheral blood lymphocytes;
RX MEDLINE=98367034; PubMed=9700200;
RA Doray B., Salomon R., Amiel J., Pelet A., Touraine R., Billaud M.,
RA Attie T., Bachy B., Munnich A., Lyonnet S.;
RT "Mutation of the RET ligand, neurturin, supports multigenic
RT inheritance in Hirschsprung disease";
RL Hum. Mol. Genet. 7:1449-1452(1998).
CC -1- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.
CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT
CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS
CC HAEMOPOIETIC CELLS.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC OTHER LOC. IT IS INVOLVED IN HIRSCHSPRUNG'S DISEASE (HSCR). THIS
CC GENETIC DISORDER OF NEURAL CREST DEVELOPMENT IS CHARACTERIZED BY
CC THE ABSENCE OF INTRAMURAL GANGLION CELLS IN THE HINDGUT; OFTEN
CC RESULTING IN INTESTINAL OBSTRUCTION.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; U78110; AAC50898.1;

DR EMBL; AL161995; CAB82327.1; -.
DR HSP; Q07731; IAGQ.
DR MIM; 602018; -.
DR MIM; 142623; -.
DR InterPro: IPR002400; GF_cysknob.
DR InterPro: IPR001839; TGF-beta.
DR PRINTS; PR00438; GFCYSKNOT.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
KW Growth factor; Signal; Disease mutation; Hirschsprung disease.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 95 BY SIMILARITY.
FT CHAIN 96 197 NEURTURIN.
FT DISULFID 103 165 BY SIMILARITY.
FT DISULFID 130 194 BY SIMILARITY.
FT DISULFID 134 196 BY SIMILARITY.
FT DISULFID 164 164 INTERCHAIN (BY SIMILARITY).
FT VARIANT 96 96 A -> S (IN HSCR; ASSOCIATED TO A RET
FT MUTATION; INCOMPLETE PENETRANCE).
FT /FTID=VAR_009498.
SQ SEQUENCE 197 AA; 22405 MW; 91AFAC8C3F8971FD CRC64;

Query Match 32.3%; Score 241; DB 1; Length 197;
Best Local Similarity 46.2%; Pred. No. 1.6e-11;
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGGAARAGGPGSRARA-AGARGCRLRSQLVPRALGLGHRSDLVRFRC 69
DB 80 PPGP-----RRRAGPRRRARALRGPCGLRELVRSVSELGLGVASDVLFRYC 130
QY 70 SGSCRRARSPHDLSLALGAGALRPPGSRPVSPQCCRPTRYE-AVSFMDVNSTWRTVD 128
DB 131 AGACEAAARYVDLCLRLRRRLR---REVRAPQCCRTAYEDEVSELDHRSYHTVH 187
QY 129 RLSATACGCL 138
DB 188 ELSARECAV 197

RESULT 3
PSPN_MOUSE STANDARD: PRT; 156 AA.
ID PSPN_MOUSE AC 070300;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PERSEPHIN PRECURSOR (PSP).
GN PSPN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=98150950; PubMed=9491986;
RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,
RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotsbauer P.T.,
RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,
RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vanden R., Moffat B.,
RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,
RA Phillips H.S., Johnson E.M.;
RT "Persephin, a novel neurotrophic factor related to GDNF and
RT neurturin";
RL Neuron 20:245-253(1998).
CC -1- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC
CC DOPAMINERGIC AND MOTOR NEURONS.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF040960; AAC40057.1; -
 DR HSP; Q07731; IAGQ.
 DR MGD; MGI:1201684; Pspn.
 DR InterPro; IPR001839; TGF-beta.
 DR SMART; SM00204; TGF-beta.
 DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
 KW Growth factor; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 156 PERSEPHIN.
 FT DISULFID 66 124 BY SIMILARITY.
 FT DISULFID 93 152 BY SIMILARITY.
 FT DISULFID 97 154 BY SIMILARITY.
 FT DISULFID 123 123 INTERCHAIN (BY SIMILARITY).
 SQ SEQUENCE 156 AA; 17030 MW; 7DC6DD98132E041B CRC64;

Query Match 31.1%; Score 232.5; DB 1; Length 156;
 Best Local Similarity 43.8%; Pred. No. 5.5e-11;
 Matches 53; Conservative 14; Mismatches 45; Indels 9; Gaps 2;

QY 20 RGGRAARAGGPGSRARAGCRLRSQVPRALGLGHRSDLVRFRCGSGC-RRARS 78
 DB 44 RGTWTHQGNHVRPLPALAGSRLWSLTPVAELGLGYASEKVFIRYFCAGSCPGERT 103
 QY 79 PHDLUSLASLIGAGALRPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCL 138
 DB 104 QHSLVLRRLRG-----RAHGRCCQPTSYADVTFELDDQHHWQLPQLSAAACGGC 155
 QY 139 G 139
 DB 156 G 156

RESULT 4
 NRTN_MOUSE
 ID NRTN_MOUSE STANDARD; PRT; 195 AA.
 AC P97463;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE NEURTURIN PRECURSOR.
 GN NRTN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 96-110; 127-135; 155-177 & 181-190.
 RX MEDLINE=97100947; PubMed=8945474;
 RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,
 RA Crendon D.J., Johnson E.M. Jr., Milbrandt J.;
 RT "Neurturin, a relative of glial-cell-line-derived neurotrophic
 factor.";
 RL Nature 384:467-470(1996).
 CC -1- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.
 CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT
 CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS
 CC HAEMOPOIETIC CELLS.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: WIDESPREAD DISTRIBUTION.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U78109; AAC52954.1; -
 DR HSP; Q07731; IAGQ.
 DR MGD; MGI:108417; Nrtin.
 DR InterPro; IPR002400; GF_CYSKNOT.
 DR InterPro; IPR001839; TGF-beta.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR SMART; SM00204; TGF-beta.
 DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
 KW Growth factor; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 95 BY SIMILARITY.
 FT CHAIN 96 195 NEURTURIN.
 FT DISULFID 101 163 BY SIMILARITY.
 FT DISULFID 128 192 BY SIMILARITY.
 FT DISULFID 132 194 BY SIMILARITY.
 FT DISULFID 162 162 INTERCHAIN (BY SIMILARITY).
 SQ SEQUENCE 195 AA; 22219 MW; ABE21BB35D417448 CRC64;

Query Match 30.9%; Score 231; DB 1; Length 195;
 Best Local Similarity 46.7%; Pred. No. 8.5e-11;
 Matches 57; Conservative 12; Mismatches 43; Indels 10; Gaps 4;

QY 24 AARAGP-----CSRARA--GARGCRLRSQVPRALGLGHRSDLVRFRCGSGCRRAR 77
 DB 77 AARIPGPRRAGRRRRRARGPCGLRELVRSVSELGYTSDETVLFVRYCAGACEAAI 136
 QY 78 SPDLUSLASLIGAGALRPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACG 136
 DB 137 RIYDGLRLRLRRVR---RERARHPCCRPATYEDVSVFLDVHSRYHTQLQELSARECA 193
 QY 137 CL 138
 DB 194 CV 195

RESULT 5

PSPN_RAT
 ID PSPN_RAT STANDARD; PRT; 156 AA.
 AC O70301;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PERSEPHIN PRECURSOR (PSP).
 GN PSPN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98150950; PubMed=9491986;
 RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,
 RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,
 RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,
 RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,
 RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,
 RA Phillips H.S., Johnson E.M.;
 RT "Persephin, a novel neurotrophic factor related to GDNF and
 neurturin.";
 RL Neuron 20:245-253(1998).
 RN [2]
 RP SEQUENCE OF 1-78 FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Pons;
 RX MEDLINE=98374044; PubMed=9710270;
 RA Jaszai J., Farkas L.M., Galtier D., Reuss B., Strelau J., Unsicker K.,
 RA Krieglstein K.;
 RT "GDNF-related factor persephin is widely distributed throughout the

RT nervous system.";
 RL J. Neurosci. Res. 53:494-501(1998).
 CC -!- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC
 CC DOPAMINERGIC AND MOTOR NEURONS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF040961; AAC40058.1; -;
 DR EMBL; AJ005169; CAA06410.1; -;
 DR HSP; Q07731; IAGQ.
 DR InterPro: IPR001839; TGF-beta.
 DR SMART: SM00204; TGFb; 1.
 DR PROSITE: PS00250; TGF_BETA_1; FALSE_NEG.
 KW Growth factor; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 156 PERSEPHIN.
 FT DISULFID 66 124 BY SIMILARITY.
 FT DISULFID 93 152 BY SIMILARITY.
 FT DISULFID 97 154 BY SIMILARITY.
 FT DISULFID 123 123 INTERCHAIN (BY SIMILARITY).
 SQ SEQUENCE 156 AA; 17063 MW; 9631941CC69B00B0 CRC64;
 Query Match 29.7%; Score 221.5; DB 1; Length 156;
 Best Local Similarity 45.5%; Pred No. 3.5e-10;
 Matches 45; Conservative 16; Mismatches 29; Indels 9; Gaps 2;
 QY 42 CLRSLQVLPVRLGLGHRSDLVRFRCGSGC-RRARSPHDLGLAGALRPPPGSR 100
 DB 66 CLRSLTLPVLAELGLYASEEIKIIFRCAGSCQPEVTRTHSLVLRGQG-----R 117
 QY 101 PVSQPCRTRYEAVSFMDVNSTWRTVRLSATACGCG 139
 DB 118 AHGRPCQPTSYADVTFELDDHHHQQLPOLSAACGCG 156
 RESULT 6
 GDNF_MOUSE
 ID GDNF_MOUSE STANDARD; PRT; 211 AA.
 AC P48340; P97919; O09058; P97920; P70446;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR PRECURSOR.
 GN GDNF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC STRAIN=C57BL/10J; TISSUE=Brain;
 RA Wang F., Too H.P.;
 RA Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=Dorsal root ganglion;
 RX MEDLINE=95379105; PubMed=7650763;
 RA Watabe K., Fukuda T., Tanaka J., Honda H., Toyohara K., Sakai O.;
 RT "Spontaneously immortalized adult mouse Schwann cells secrete
 RT autocrine and paracrine growth-promoting activities.";
 RL J. Neurosci. Res. 41:279-290(1995).
 RN [3]
 RP SEQUENCE FROM N.A.

RC STRAIN=129/SVJ;
 RA Hellmich H., Kos L., Cho E.S., Mahon K.A., Zimmer A.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Matsushita N., Fujita Y., Nagatsu T., Kiuchi K.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND
 CC MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND
 CC INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U37459; AAB18672.1; ALT_INIT.
 DR EMBL; U66195; AAB07463.1; ALT_INIT.
 DR EMBL; U75532; AAB18343.1; ALT_INIT.
 DR EMBL; D49921; BAA08660.1; -;
 DR EMBL; U36449; AAB52953.1; -;
 DR EMBL; D88264; BAA13566.1; ALT_INIT.
 DR EMBL; D88352; BAB12221.1; -;
 DR HSP; Q07731; IAGQ.
 DR MGD; MGI:107430; Gdnf.
 DR InterPro: IPR001839; TGF-beta.
 DR SMART: SM00204; TGFb; 1.
 DR PROSITE: PS00250; TGF_BETA_1; FALSE_NEG.
 KW Growth factor; Glycoprotein; Signal; Alternative splicing.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 77 BY SIMILARITY.
 FT CHAIN 78 211 GLIAL CELL LINE-DERIVED NEUROTROPHIC
 FT FACTOR.
 FT DISULFID 118 179 BY SIMILARITY.
 FT DISULFID 145 208 BY SIMILARITY.
 FT DISULFID 149 210 BY SIMILARITY.
 FT DISULFID 178 178 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 25 51 GKRLLEAPAEHSLGHRVFFALTSDS -> A (IN
 FT ISOFORM 2).
 SQ SEQUENCE 211 AA; 23662 MW; B6731C767A3A95B7 CRC64;
 Query Match 25.1%; Score 187.5; DB 1; Length 211;
 Best Local Similarity 35.0%; Pred No. 1.4e-07;
 Matches 48; Conservative 21; Mismatches 55; Indels 13; Gaps 3;
 QY 11 PPAPPSALPRGGRARAG-----GPGSRARAAGARGCRLRSQVLPVRLGLGHRSD 62
 DB 79 PDQAAALPRRERNRQAASAPENSRGKRGQKRGCVLTAHLNVITDLGLGYETKE 138
 QY 63 LVRFRCGSGCRRARSPHDLGLAGALRPPPGSRVPSQPCRTRY-EAVSFMDVN 121
 DB 139 ELIFRYCGSGCESAETMYDKILKNLSRRLT-----SDKVGQACCRPVAFDDDLSD 194
 QY 122 STWRTVRLSATACGCG 138
 DB 195 LVYHILRKHSAKRCGI 211
 RESULT 7
 GDNF_RAT
 ID GDNF_RAT STANDARD; PRT; 211 AA.

neurotrophic factor gene.";
 Brain Res. Mol. Brain Res. 69:209-222(1999).
 [4]
 PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 MEDLINE-97141760; PubMed-8988018;
 Hanlu M., Hui J., Young Y., Le J., Katta V., Lee R., Shimamoto G.,
 Rohde M.F.;
 "Glial cell line-derived neurotrophic factor: selective reduction of
 the intermolecular disulfide linkage and characterization of its
 disulfide structure";
 Biochemistry 35:16799-16805(1996).
 [5]
 REVIEW ON VARIANTS.
 MEDLINE-98023959; PubMed-9359036;
 Hofstra R.M.W., Ozinga J., Buys C.H.C.M.;
 "Mutations in Hirschsprung disease: when does a mutation contribute to
 the phenotype.";
 Eur. J. Hum. Genet. 5:180-185(1997).
 [6]
 VARIANT HSCR SER-154.
 MEDLINE-97123511; PubMed-8968758;
 Ivanchuk S.M., Myers S.M., Eng C., Mulligan L.M.;
 "De novo mutation of GDNF, ligand for the RET/GDNF-alpha receptor
 complex, in Hirschsprung disease.";
 Hum. Mol. Genet. 5:2023-2026(1996).
 [7]
 VARIANT HSCR TRP-93.
 MEDLINE-97051933; PubMed-8996568;
 Angrist M., Bolk S., Haluska M., Lapchak P.A., Chakravarti A.;
 "Germline mutations in glial cell line-derived neurotrophic factor
 (GDNF) and RET in a Hirschsprung disease patient.";
 Nat. Genet. 14:341-344(1996).
 [8]
 VARIANTS HSCR SER-21 AND ASN-150.
 MEDLINE-97051934; PubMed-8996569;
 Salomon R., Attie T., Pelet A., Bidaud C., Eng C., Amiel J.,
 Sarnacki S., Goulet O., Ricour C., Nihoul-Fekete C., Munnich A.,
 Lyonnet S.;
 "Germline mutations of the RET ligand GDNF are not sufficient to cause
 Hirschsprung disease.";
 Nat. Genet. 14:345-347(1996).
 [9]
 FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND
 MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND
 INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.
 [10]
 SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 [11]
 SUBCELLULAR LOCATION: SECRETED.
 [12]
 DISEASE: IN ASSOCIATION WITH MUTATIONS OF RET GENE, MAY BE
 INVOLVED IN HIRSCHSPRUNG'S DISEASE (HSCR). THIS GENETIC DISORDER
 OF NEURAL CREST DEVELOPMENT IS CHARACTERIZED BY THE ABSENCE OF
 INTRAMURAL GANGLION CELLS IN THE HINDGUT; OFTEN RESULTING IN
 INTESTINAL OBSTRUCTION.
 [13]
 ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 PRODUCED BY ALTERNATIVE SPLICING.
 [14]
 SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 [15]
 DATABASE: NAME=RD Systems' cytokine source book;
 WWW="http://www.rndsystems.com/cyt_cat/gdnf.html".
 [16]
 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 [17]
 EMBL: L19063; AAA67910.1;
 EMBL: L19062; AAA67910.1; JOINED.
 EMBL: AF053748; AAD43139.1;
 PIR: B37499; B37499.
 HSP: Q07731; LAGO.
 MIM: 600837;
 MIM: 142623;
 InterPro: IPR001839; TGF-beta.

SMART; SM00204; TGFβ; 1.
 PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
 Growth factor; Glycoprotein; Signal; Alternative splicing;
 Polymorphism; Disease mutation; Hirschsprung disease.
 SIGNAL 1 19 POTENTIAL.
 PROPEP 20 77 BY SIMILARITY.
 CHAIN 78 211 GLIAL CELL LINE-DERIVED NEUROTROPHIC
 FACTOR.
 DISULFID 118 179
 DISULFID 145 208
 DISULFID 149 210
 DISULFID 178 178
 CARBOHYD 126 126
 CARBOHYD 162 162
 VARSPIC 25 51
 VARIANT 21 21
 VARIANT 93 93
 VARIANT 150 150
 VARIANT 154 154
 SEQUENCE 211 AA; 23720 MW; A0D1EBF77FC82691 CRC64;
 Query Match 23.5%; Score 175.5; DB 1; Length 211;
 Best Local Similarity 34.3%; Pred. No. 1e-06;
 Matches 47; Conservative 20; Mismatches 57; Indels 13; Gaps 3;
 QY 11 PPAPPSALPCGGRARAG-----CPGSRARAAGARGCLRSQSLVPVPRALGLGHRSD 62
 DB 79 PDKOMAVLPRRRRNRQAAANPENSRRGRRGRRGKRGCVLTALHNVTDLGLGYETKE 138
 QY 63 LVRRFCGSGCRRARSPhDLASLALGAGALRPPPSRPSQPCRPTRY-EAVSFMDVN 121
 DB 139 ELIFRYCGSGCDRAETTYDKLNLSNRRL---VSDKVGQACCRPIAFDDDLDFLDN 194
 QY 122 STWRTVRLSATAGCL 138
 DB 195 LVYHILRKHSAKRCGI 211
 RESULT 9
 GDF6_BOVIN
 ID GDF6_BOVIN STANDARD; PRT; 436 AA.
 AC P55106;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE GROWTH/DIFFERENTIATION FACTOR 6 PRECURSOR (GDF-6) (CARTILAGE-DERIVED
 DE MORPHOGENETIC PROTEIN 2) (CDMP-2) (FRAGMENT).
 GN GDF6 OR CDMP2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Articular cartilage;
 RA MEDLINE=95050604; PubMed=7961761;
 RA Chang S., Hoang B., Thomas J.T., Vukicevic S., Luyten F.P.,
 Ryba N.J.P., Kozak C.A., Reddi A.H., Moos M.;
 RT "Cartilage-derived morphogenetic proteins. New members of the
 RT transforming growth factor-beta superfamily predominantly expressed
 RT in long bones during human embryonic development.";
 RL J. Biol. Chem. 269:28227-28234(1994).

RP SEQUENCE FROM N.A.
 RX MEDLINE=86218082; PubMed=3754790;
 RA Cate R.L., Mattaliano R.J., Hession C., Tizard R., Farber N.M., Fisher R.A.,
 RA Cheung A., Ninfa E.G., Frey A.Z., Gash D.J., Chow E.P., Fisher R.A.,
 RA Bertoni J.M., Torres G., Wallner B.P., Ramachandran K.L.,
 RA Ragin R.C., Manganaro T.F., McLaughlin D.T., Donahoe P.K.;
 RT "Isolation of the bovine and human genes for Mullerian inhibiting
 substance and expression of the human gene in animal cells.";
 RL Cell 45:685-698(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
 RA Burkhardt-Schultz K.J., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Do L., Regalia W., Terry A., Ganes J.,
 RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
 RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,
 RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
 RA Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
 RA Kobayashi A., Olsen A.S., Carrano A.V.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP VARIANT ARG-325.
 RX MEDLINE=93131268; PubMed=1483695;
 RA Carre-Eusebe D., Imbeaud S., Harbison M., New M.I., Josso N.,
 RA Picard J.Y.;
 RT "Variants of the anti-Mullerian hormone gene in a compound
 heterozygote with the persistent Mullerian duct syndrome and his
 family.";
 RT Hum. Genet. 90:389-394(1992).
 RN [4]
 RP VARIANTS PMDS-1 G-12; P-70; V-101; W-123; C-167; C-194 AND A-477.
 RX MEDLINE=94214429; PubMed=8162013;
 RA Imbeaud S., Carre-Eusebe D., Rey R., Belville C., Josso N.,
 RA Picard J.Y.;
 RT "Molecular genetics of the persistent mullerian duct syndrome: a
 study of 19 families.";
 RT Hum. Mol. Genet. 3:125-131(1994).
 CC -!- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
 CC DUCT ORIGIN.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -!- DISEASE: DEFECTS IN AMH ARE THE CAUSE OF PERSISTENT MUELLERIAN
 CC DUCT SYNDROME TYPE I (PMDS-1); A FORM OF MALE
 CC PSEUDOHERMAPHRODITISM CHARACTERIZED BY A FAILURE OF MUELLERIAN
 CC DUCT REGRESSION IN OTHERWISE NORMAL MALES.
 CC -!- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR
 CC BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF
 CC RECEPTOR IN VITRO.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; K03474; AAC98805.1; -;
 DR EMBL; AC005263; AAC25614.1; -;
 DR PIR; A01397; WFLUM.
 DR MIM; 600957; -;
 DR MIM; 261550; -;
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGF-beta.
 DR Pfam; PF00019; TGF-beta; 1.
 DR PRINTS; PR00438; GFCYSNOT.
 DR ProDom; PD000357; -; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA.1; 1.
 KW Growth factor; Glycoprotein; Gonadal differentiation; Signal;
 Pseudohermaphroditism; Disease mutation; Polymorphism.

FT SIGNAL 1 18
 FT PROPEP 19 25
 FT CHAIN 26 560
 FT DISULFID 462 526
 FT DISULFID 488 557
 FT DISULFID 492 559
 FT DISULFID 525 525
 FT CARBOHYD 64 64
 FT CARBOHYD 329 329
 FT VARIANT 12 12
 FT VARIANT 49 49
 FT VARIANT 70 70
 FT VARIANT 101 101
 FT VARIANT 123 123
 FT VARIANT 167 167
 FT VARIANT 185 185
 FT VARIANT 194 194
 FT VARIANT 325 325
 FT VARIANT 477 477
 FT CONFLICT 515 515
 FT SEQUENCE 560 AA; 59192 MW; 3EFC2E4FECC364C CRC64;
 SQ
 Query Match 14.8%; Score 110.5; DB 1; Length 560;
 Best Local Similarity 27.4%; Pred. No. 0.14;
 Matches 49; Conservative 11; Mismatches 62; Indels 57; Gaps 9;
 QY 5 SRPAPPAPP-----SALPRGG-----RAARAGGPGSRAR 35
 DB 392 SLPLPPATAPLLARLLALCPGGPGGLGDPRLALLLLKALQGLRVEWRGDRPGRAOR 451
 QY 36 AACARG-----CRLRSQLVPRVRLGLGHR-----DELVRFRFCGSC-----RRARSP-----HD 81
 DB 452 SAGATAADGACALRELSVDLRA-----ERSVLIPETQANNCGVCGWQSDRNPYGNHV 507
 QY 82 LSLASLLGACALRPPGCSRPVOPCCRPTRYEA---VSEFMDVNSTWRTVDRLSATACGC 137
 DB 508 VLLKMQARGAALARP-----PCCVPTAYAGKLLISLEERISAHHPVNVWATECGC 559
 RESULT 12
 MIS_RAT STANDARD; PRT; 553 AA.
 ID MIS_RAT
 AC P49000;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)
 DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).
 GN AMH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92241861; PubMed=1572639;
 RA Haqq C., Lee M.M., Tizard R., Wysk M., Demarinis J., Donahoe P.K.,
 RA Cate R.L.;
 RT "Isolation of the rat gene for Mullerian inhibiting substance.";
 RL Genomics 12:665-669(1992).
 CC -!- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE

TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN DUCT ORIGIN.

-1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.

-1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: S98336; AAB22104.1; .

InterPro: IPR002400; GE-CysKnot.

InterPro: IPR001839; TGF-beta.

Pfam: PF00019; TGF-beta; 1.

PRINTS: PR00438; GFCYSKNOT.

ProDom: PD000357; .; 1.

SMART: SM00204; TGF-beta; 1.

PROSITE: PS00250; TGF_BETA_1; 1.

Growth factor; Glycoprotein; Gonadal differentiation; Signal.

SIGNAL 1 21

FT PROPEP 22 ? POTENTIAL.

FT CHAIN 2 ? MUELLERIAN INHIBITING FACTOR.

FT DISULFID 455 519 BY SIMILARITY.

FT DISULFID 481 550 BY SIMILARITY.

FT DISULFID 485 552 BY SIMILARITY.

FT DISULFID 518 518 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 325 325 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 409 409 N-LINKED (GLCNAC. .) (POTENTIAL).

SEQUENCE 553 AA; 58888 MW; 75DAF3949A038A69 CRC64;

Query Match 14.3%; Score 107; DB 1; Length 553;

Best Local Similarity 26.4%; Pred. No. 0.24;

Matches 46; Conservative 16; Mismatches 58; Indels 54; Gaps 10;

QY 7 PAPPAPP-----SALPRGGRA-----RAG-GPGSRARAA 37

DB 390 PGLPTAPPALLSRLLALCPNDSRSAGPLRALLLLKALQGLRAEWREGGRGRAGRSKGT 449

QY 38 GARG-CRLRSQVPPVRLGLHRS---DELVRFRCGSC---RRASP-----HDLSLAS 86

DB 450 GTDGLCALRELSVDLRA-----ERSVLIPETQYANNCQACAWPQSDRNPRYGNHVLLLK 505

QY 87 LLGAGALRPPPGSRVSPQCCPRTRYEA---VSFMDVNSTWRTVDRLSATACGC 137

DB 506 MQARGAALG-----RLPCCVPTATGKLLISLSEEHISAHVPMVATECGC 552

RESULT 13

GDFP_RAT

ID GDFP_RAT STANDARD; PRT; 303 AA.

AC Q92076;

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE GROWTH/DIFFERENTIATION FACTOR 15 PRECURSOR (GDF-15).

GN GDF15 OR SBF.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Boettner M., Laaff M., Suter-Crazzolara C.;

RT "Identification of a novel member of the TGFbeta superfamily."

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AJ011969; CAA09891.1; .

EMBL: AJ011970; CAA09891.1; JOINED.

HSSP: P18075; IBMP.

InterPro: IPR001839; TGF-beta.

Pfam: PF00019; TGF-beta; 1.

ProDom: PD000357; .; 1.

SMART: SM00204; TGF-beta; 1.

PROSITE: PS00250; TGF_BETA; FALSE NEG.

Growth factor; Cytokine; Glycoprotein; Signal.

SIGNAL 1 30

FT PROPEP 31 188 POTENTIAL.

FT CHAIN 189 303 GROWTH/DIFFERENTIATION FACTOR 15.

FT DISULFID 206 269 BY SIMILARITY.

FT DISULFID 235 300 BY SIMILARITY.

FT DISULFID 239 302 BY SIMILARITY.

FT DISULFID 268 268 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 71 71 N-LINKED (GLCNAC. .) (POTENTIAL).

SEQUENCE 303 AA; 33438 MW; FA828BE79C1052C9 CRC64;

Query Match 14.1%; Score 105.5; DB 1; Length 303;

Best Local Similarity 29.2%; Pred. No. 0.19;

Matches 45; Conservative 12; Mismatches 66; Indels 31; Gaps 8;

QY 6 RPAPPPAPPAPPALPRGG-----RAARAGG-----PGSRARAGAGCRLRSQVLPV 51

DB 158 RLAPPPDL--AVLPGGARLEHLHLSAAGRRSAHLHPRDSCPLGRCRCHLETVOATL 215

QY 52 RALGLGH--RSDELVRFRCGSC---RRARSPHDLSLALLGAGALRPPPGSRVSPQC 106

DB 216 EDLGSWDVLSPROLQSLSMCVGCPHLYRSANTHALIKARLHG---LQPD-----RVAPAC 268

QY 107 CRPRTRYEAIVSPM---DYNSTWRTVDRLSATACGC 137

DB 269 CVPSSTPVVLMHRTDGSVSLQTYDDLVAQGCCHC 302

RESULT 14

GDFP_MOUSE

ID GDFP_MOUSE STANDARD; PRT; 303 AA.

AC Q92007;

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE GROWTH/DIFFERENTIATION FACTOR 15 PRECURSOR (GDF-15).

GN GDF15 OR SBF.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=129/SVJ;

RA Boettner M., Laaff M., Suter-Crazzolara C.;

RT "Identification of a novel member of the TGFbeta superfamily."

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Hsiao E.C., Koniaris L.G., Zimmers T.A., Sebald S.M., Sitzmann J.V.,

RA Huynh T.V., Lee S.-J.;

RT "Growth/differentiation factor-15: a new TGF-beta family member induced following liver and bile duct injury."

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: SECRETED (PROBABLE).

```

CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ011967; CAA09890.1; -.
DR EMBL: AJ011968; CAA09890.1; JOINED.
DR EMBL: AF159571; AAD41410.1; -.
DR HSSP: P18075; 1BMP.
DR MGD: MGI:1346047; Gdf15.
DR InterPro: IPR001839; TGF-beta.
DR Pfam: PF00019; TGF-beta; 1.
DR ProDom: PD000357; -. 1.
DR SMART: SM00204; TGF-beta; 1.
DR PROSITE: PS00250; TGF_BETA; FALSE_NEG.
KW Growth factor; Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 30
FT PROPEP 31 188
FT CHAIN 189 303
FT DISULFID 206 269
FT DISULFID 235 300
FT DISULFID 239 302
FT DISULFID 268 268
FT CARBOHYD 71 71
SQ SEQUENCE 303 AA; 33256 MW; C7AB5F3CBE5639B9 CRC64;

Query Match 13.9%; Score 103.5; DB 1; Length 303;
Best Local Similarity 29.0%; Pred. No. 0.26; Mismatches 66; Indels 37; Gaps 9;
Matches 47; Conservative 12;

QY 4 PSRPA-----PPPPAPPSALPRGG-----RAARAGSGPSRARAGARG-----CR 43
Db 150 PRAPALRLRLTPPDLL--AMLPSSGGAQLRLRLVAAGRRSAHAHPDSCPLGPGRCCH 207
QY 44 LRSQLVPRALGLGH--RSDELVRFRCGSC--RRARSPHDLASLLAGALRPPPG 98
Db 208 LFTVQATLEDLGMWDLVSPQLQLSMCVGCPHLYRSANTHAQIKARLHG---LQPD-- 262
QY 99 SRPVSOPCCRPTRYEAVSFM---DYNSTWRTVDRLSATACGC 137
Db 263 --KVPAFCVPSSTYFVVLHMTDGSVSLQTYDDLVARGCCH 302

RESULT 15
GDFP_HUMAN STANDARD; PRT; 308 AA.
AC Q99588; P78360; O14629; Q9NRT0;
DT 01-OCT-2000 (Rel. 40; Created)
DT 01-OCT-2000 (Rel. 40; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE GROWTH/DIFFERENTIATION FACTOR 15 PRECURSOR (GDF-15) (PLACENTAL BONE
DE MORPHOGENIC PROTEIN) (PLACENTAL TGF-BETA) (MACROPHAGE INHIBITORY
DE CYTOKINE-1) (MIC-1) (PROSTATE DIFFERENTIATION FACTOR) (NSAID-REGULATED
DE PROTEIN 1) (NRG-1).
GN GDF15 OR PLAB OR PTGFB OR MIC1 OR PDF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblasts;
RX MEDLINE=98006316; PubMed=9348093;
RA Yokoyama-Kobayashi M., Saeki M., Sekine S., Kato S.;
RT "Human cDNA encoding a novel TGF-beta superfamily protein highly
RT expressed in placenta.";
RL J. Biochem. 122:622-626(1997).

```

```

RN RP SEQUENCE FROM N.A.
RX MEDLINE=97470998; PubMed=9326641;
RA Bootcov M.R., Bauskin A.R., Valenzuela S.M., Moore A.G., Bansal M.,
RA He X.Y., Zhang H.P., Donnellan M., Mahler S., Pryor K., Walsh B.J.,
RA Nicholson R.C., Fairlie W.D., Por S.B., Robbins J.M., Brett S.N.;
RT "MIC-1, a novel macrophage inhibitory cytokine, is a divergent member
RT of the TGF-beta superfamily.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11514-11519(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=98041637; PubMed=9375789;
RA Hromas R., Hufford M., Sutton J., Xu D., Li Y., Lu L.;
RT "PLAB, a novel placental bone morphogenetic protein.";
RL Biochim. Biophys. Acta 1354:40-44(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=98256302; PubMed=9593718;
RA Paralkar V.M., Vail A.L., Grasser W.A., Brown T.A., Xu H.,
RA Vukicevic S., Ke H.Z., Qi H., Owen T.A., Thompson D.D.;
RT "Cloning and characterization of a novel member of the transforming
RT growth factor-beta/bone morphogenetic protein family.";
RL J. Biol. Chem. 273:13760-13767(1998).
RN [5]
RP SEQUENCE OF 14-308 FROM N.A.
RX MEDLINE=98085971; PubMed=9426002;
RA Lawton L.N., de Fatima Bonaldo M., Jelenc P.C., Oiu L., Baumes S.A.,
RA Marcelino R.A., de Jesus G.M., Wellington S., Knowles J.A.,
RA Warburton D., Brown S., Soares M.B.;
RT "Identification of a novel member of the TGF-beta superfamily highly
RT expressed in human placenta.";
RL Gene 203:17-26(1997).
RN [6]
RP SEQUENCE OF 264-308 FROM N.A.
RA Baek S.J., Nixon J., Eling T.;
RT "NRG-1 is associated with apoptosis.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA, WITH LOWER
CC LEVELS IN PROSTATE AND COLON AND SOME EXPRESSION IN KIDNEY.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB000584; BAAL19151.1; -.
DR EMBL: AF019770; AAB88673.1; -.
DR EMBL: U88323; AAB88913.1; -.
DR EMBL: AF003934; AAC24456.1; -.
DR EMBL: AF008303; AAC39537.1; -.
DR EMBL: AF173860; AAF89834.1; -.
DR MIM: 605312; -.
DR HSSP: P18075; 1BMP.
DR InterPro: IPR002400; GF_Cyskn0t.
DR InterPro: IPR001839; TGF-beta.
DR Pfam: PF00019; TGF-beta; 1.
DR PRINTS: PR00438; GFCYSKN0T.
DR ProDom: PD000357; -. 1.
DR SMART: SM00204; TGF-beta; 1.
DR PROSITE: PS00250; TGF_BETA; FALSE_NEG.
KW Growth factor; Cytokine; Glycoprotein; Signal; Polymorphism.
FT SIGNAL 1 29
FT PROPEP 30 194
FT CHAIN 195 308
FT DISULFID 211 274

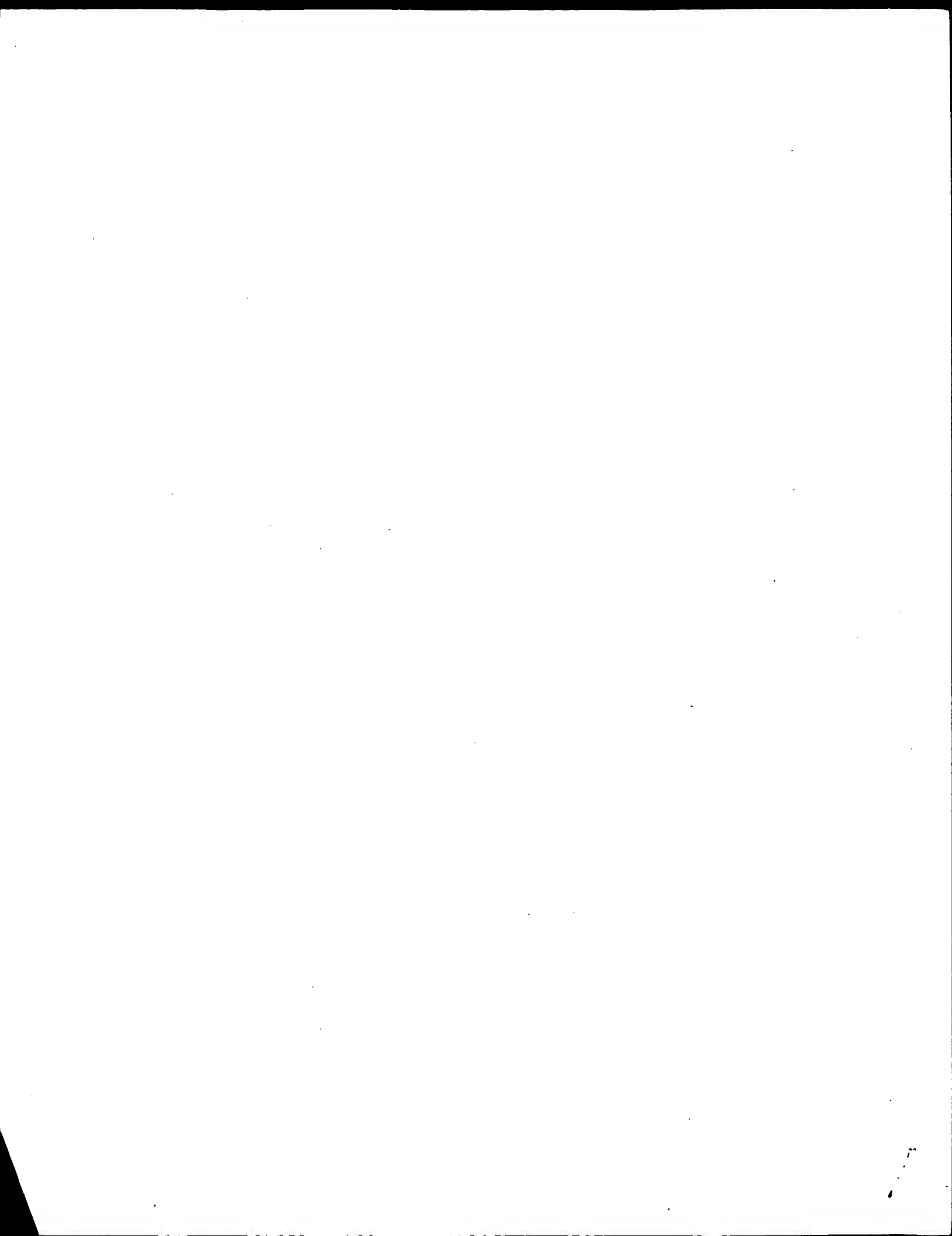
```

```
FT DISULFID 240 305 BY SIMILARITY.
FT DISULFID 244 307 BY SIMILARITY.
FT DISULFID 273 273 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 48 48 T -> S.
FT          /FTID=VAR_010386.
FT CONFLICT 9 9 L -> V (IN REF. 3).
FT CONFLICT 202 202 H -> D (IN REF. 3).
FT CONFLICT 269 269 V -> E (IN REF. 3).
FT CONFLICT 288 288 T -> A (IN REF. 6).
SQ SEQUENCE 308 AA; 34168 MW; A0F3A3ED065ACA2E CRC64;
```

Query Match 13.7%; Score 102; DB 1; Length 308;
Best Local Similarity 28.0%; Pred. No. 0.34;
Matches 49; Conservative 13; Mismatches 59; Indels 54; Gaps 10;

```
QY 3 QPSRPA-----PPPA-----PPSALP-----RGGRARAG-----GPG 31
Db 149 RQAPALHURLSPPSQDQLAESSARPQLEHLRPOAARGRRARNGDHCPLGPG 208
QY 32 SRARAAGRCRLRSQVPRALGLGH--RSDELVRFRCGSCR---RARSPHDLSLAS 86
Db 209 -----RCCRLHTVRASLEDLGHADWVLSPREVQVTCIGACPSQFRAANMHAQIKTS 260
QY 87 LLGAGALRPPGSRPVSQPCRPTRYEAVSFM---DVNSTWRTVDRLSATACGCL 138
Db 261 L----HRLKPD----TVPAPCCVPASYNPMVLQKTDGTGVSQTYDDLLAKDCHCI 308
```

Search completed: August 17, 2001, 09:54:44
Job time: 24 sec



Query: Local Similarity 34.3%; Score 179.3; E-Value 0.002; Length 211; Best Local Similarity 34.3%; Pred. No. 4.1e-07; Matches 47; Conservative 20; Mismatches 57; Indels 13; Gaps 3;

| | | |
|----|-----|---|
| Qy | 11 | PPAPSPALPRGGRAARG-----GPGSRARAAGARGCRLRSQLVPRALGLGHRSD 62 |
| | | : |
| Db | 79 | PDQMAYLPRERNRQAAANPENSRCGRGRGQGNRGCVLTALHINVTDLGLGYETKE 138 |
| | | : |
| Qy | 63 | LVRPFPGSCRRARSHPDLSLASLLGAGALRPPGSGRPVSPCCRPTRY-EAVSPMDVN 121 |
| | | : |
| Db | 139 | ELIFRYGSCDAAETTYDKILNLSNRRL---VSDKVGQACCRPIAFDDDLSDLDN 194 |
| | | : |
| Qy | 122 | STWRTVDRLSATCGCL 138 |
| | | : |
| Db | 195 | LVYHILRKHSAKRCCCI 211 |
| | | : |

RESULT 5
B55452
cartilage-derived morphogenetic protein 2 precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 26-May-2000
C:Accession: B55452
R:Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; J. Biol. Chem. 269, 28227-28234, 1994

Query Match 13.7%; Score 102; DB 2; Length 309;
Best Local Similarity 28.0%; Pred. No. 0.32;
Matches 49; Conservative 13; Mismatches 59; Indels 54; Gaps 10;

QY 3 QPSRPA-----PPPPA-----PPSALP-----RGGRAARAG-----GPG 31
 DB 150 RQAPALHLRLSPPSQDQLAESSARPQLEHLRFPQARRRRARRNGDHCPGPG 209
 QY 32 SRRAAGARGCRLRSOLVPRALGLGH--RSDELVRFRFCGSGCR---RARSPhdLSLAS 86
 DB 210 -----RCRLHVTVRASLEDLGHADWVLSPREVOVTWICGACFSQFRAANWHAQIKTS 261
 QY 87 LLGAGALRPPGSRPVSPQCCRPTRYEAVSFM---DVNSTWRTVDRLSATACGCL 138
 DB 262 L---HRLKPD---TVPAPCCVPSYNPVMVLIQKTDGTGVSLSQTYDDLLAKDCHCI 309

RESULT 13
 B45344
 probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)
 C:Species: suid herpesvirus 1
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
 C:Accession: B45344
 R:Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwytzer, M.
 Virology 179, 365-377, 1990
 A:Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented or
 A:Reference number: A45344; MUID:91021039
 A:Accession: B45344
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1733 <VLC>
 A:Cross-references: GB:M34651; NID:g334070; PIDN:AAA47471.1; PID:g334072
 C:Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 13.7%; Score 102; DB 1; Length 1733;
 Best Local Similarity 46.0%; Pred. No. 1.4;
 Matches 23; Conservative 0; Mismatches 15; Indels 12; Gaps 2;
 QY 1 PPQSRPAPPAPPSPALPRGG---RAARAGCP-----GSRARAAG 38
 DB 280 PPPPLPPPPPPPPPPPPAGGSARRRRGGGPPGRRGGRRRRAEG 329

RESULT 14
 S20100
 mullerian inhibiting factor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Dec-2000
 C:Accession: S20100; S51159
 R:Muensterberg, A.; Lovell-Badge, R.
 Development 113, 613-624, 1991
 A:Title: Expression of the mouse anti-Muellerian hormone gene suggests a role in both ma
 A:Reference number: S20100; MUID:92146272
 A:Accession: S20100
 A:Molecule type: DNA
 A:Residues: 1-555 <MUE>
 A:Cross-references: EMBL:X63240; NID:g49945; PIDN:CAA44912.1; PID:g49946
 R:Dresser, D.W.; Hacker, A.; Lovell-Badge, R.; Guerrier, D.
 submitted to the EMBL Data Library, January 1995
 A:Description: The genes for anti-Muellerian hormone (AMH) and a spliceosome protein (SA
 A:Reference number: S51159
 A:Accession: S51159
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-41 <DRE>
 A:Cross-references: EMBL:X83733
 C:Genetics:
 A:Introns: 135/1; 182/3; 219/1; 272/2
 C:Superfamily: inhibin

Query Match 13.6%; Score 101.5; DB 1; Length 555;
 Best Local Similarity 26.3%; Pred. No. 0.58;
 Matches 46; Conservative 12; Mismatches 62; Indels 55; Gaps 9;
 QY 7 PAPPPPAPP-----SALPRGGA-----ARAGGPGSRARAAGARG- 41

DB 391 PGLPPTAPPLARLLALCPNDSSSGDPLRALLLKALQGLRAEWHGREGRTAQRGD 450
 QY 42 -----CRLRSQLVPRALGLGHR-----DELVRFRFCGSGCR---RARSPhdLSLA 85
 DB 451 KGQDPCALRELSVDLRA-----ERSVLIPETYQANNCQACRWQSDRNPRYGNHVVLL 506
 QY 86 SLLGAGALRPPGSRPVSPQCCRPTRYEA---VSFMDVNSTWRTVDRLSATACGC 137
 DB 507 KMQARGAALG-----RLPCCVPTAVAGKLLISLSEERISADHVPNMVATECGC 554

RESULT 15
 B40505
 hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
 C:Species: suid herpesvirus 1
 C:Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 02-Sep-2000
 C:Accession: B40505
 R:Cheung, A.K.
 J. Virol. 65, 5260-5271, 1991
 A:Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies vir
 A:Reference number: A40505; MUID:91374576
 A:Accession: B40505
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1958 <CHE>
 A:Cross-references: GB:M57505; NID:g334066; PIDN:AAA47468.1; PID:g334068
 C:Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 13.6%; Score 101.5; DB 2; Length 1958;
 Best Local Similarity 45.8%; Pred. No. 1.6;
 Matches 22; Conservative 1; Mismatches 22; Indels 3; Gaps 1;
 QY 1 PPQSRPAPPAPPSPALPRGG---RAARAGGPGSRARAAGRCRLR 45
 DB 489 PPPPLPPPPPPPPPPPPAGGSARRRRGGGPPGRRGGRRRGGKRR 536

Search completed: August 16, 2001, 15:40:06
 Job time: 96 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:49:24 ; Search time 35.16 Seconds
(without alignments)
523.049 Million cell updates/sec

Title: US-09-357-349-4
Perfect score: 139
Sequence: 1 PPQSPAPPAPPAPPALPR.....VNSTRVTRLSATACGCLG 139

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTRMBL16.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_unclassified.*
- 13: sp_vertebrate.*
- 14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 139 | 100.0 | 220 | 4 | O96030 homo sapien |
| 2 | 139 | 100.0 | 237 | 4 | O95441 homo sapien |
| 3 | 30 | 21.6 | 125 | 11 | O9QZG3 |
| 4 | 26 | 18.7 | 224 | 11 | O9ZOL2 |
| 5 | 10 | 7.2 | 815 | 11 | O88839 |
| 6 | 10 | 7.2 | 816 | 11 | O9QYV0 |
| 7 | 9 | 6.5 | 113 | 10 | O80640 |
| 8 | 9 | 6.5 | 136 | 14 | O9WNX1 |
| 9 | 9 | 6.5 | 137 | 2 | O05284 |
| 10 | 9 | 6.5 | 202 | 14 | O82444 |
| 11 | 9 | 6.5 | 206 | 2 | O52152 |
| 12 | 9 | 6.5 | 207 | 2 | O9RQEI |
| 13 | 9 | 6.5 | 248 | 2 | O85622 |
| 14 | 9 | 6.5 | 297 | 2 | O9RD48 |
| 15 | 9 | 6.5 | 428 | 14 | O56229 |
| 16 | 9 | 6.5 | 429 | 14 | O56621 |
| 17 | 9 | 6.5 | 429 | 14 | O82322 |
| 18 | 9 | 6.5 | 429 | 14 | O82230 |
| 19 | 9 | 6.5 | 449 | 14 | O9QRA0 |

| | | | | | |
|----|---|-----|------|----|--------|
| 20 | 9 | 6.5 | 647 | 14 | O9QR99 |
| 21 | 9 | 6.5 | 650 | 14 | O56227 |
| 22 | 9 | 6.5 | 763 | 2 | O9XDH2 |
| 23 | 9 | 6.5 | 814 | 4 | O13444 |
| 24 | 9 | 6.5 | 814 | 4 | O13493 |
| 25 | 9 | 6.5 | 1273 | 14 | O9QRA1 |
| 26 | 9 | 6.5 | 1461 | 14 | O56228 |
| 27 | 8 | 5.8 | 70 | 6 | O9XSR8 |
| 28 | 8 | 5.8 | 75 | 5 | O62468 |
| 29 | 8 | 5.8 | 75 | 5 | O9U3L6 |
| 30 | 8 | 5.8 | 141 | 13 | O9PV28 |
| 31 | 8 | 5.8 | 211 | 10 | O9M644 |
| 32 | 8 | 5.8 | 228 | 2 | O9L129 |
| 33 | 8 | 5.8 | 378 | 2 | O53236 |
| 34 | 8 | 5.8 | 393 | 5 | O44163 |
| 35 | 8 | 5.8 | 395 | 2 | O9L124 |
| 36 | 8 | 5.8 | 501 | 10 | O9SP93 |
| 37 | 8 | 5.8 | 509 | 5 | O9GS27 |
| 38 | 8 | 5.8 | 622 | 11 | O9EQM1 |
| 39 | 8 | 5.8 | 647 | 5 | O9VS45 |
| 40 | 8 | 5.8 | 678 | 2 | P71707 |
| 41 | 8 | 5.8 | 810 | 2 | O69650 |
| 42 | 8 | 5.8 | 825 | 11 | O9QXW6 |
| 43 | 8 | 5.8 | 908 | 11 | O9R037 |
| 44 | 8 | 5.8 | 912 | 6 | O9XSC3 |
| 45 | 8 | 5.8 | 923 | 5 | O9VAG7 |

ALIGNMENTS

RESULT 1
O96030 PRELIMINARY; PRT; 220 AA.
AC O96030;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ARTEMIN.
GN ARTN OR EVN.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95098192; PubMed=9883723;
RA Balch R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,
RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,
RA Milbrandt J.;

RT "Artemin, a novel member of the GDNF ligand family, supports peripheral and central neurons and signals through the GFRalpha3-RET receptor complex.";
RL Neuron 21:1291-1302(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=BRAIN;
RA Hansen C., Blom N., Johansen T.E., GDNF ligand family.";
RT "Neublastin a novel member of the GDNF ligand family.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20050601; PubMed=10583383;
RA Masure S., Geerts H., Cik M., Hoefnagel E., Van Den Kieboom G.,
RA Tuytelaars A., Harris S., Lesage A.S., Leysen J.E., van der Helm L.,
RA Verhaesselt P., Von J., Gordon R.D.;

RT "Enovin, a member of the glial cell-line-derived neurotrophic factor (GDNF) family with growth-promoting activity on neuronal cells.
RT Existence and tissue-specific expression of different splice variants.";
RL Eur. J. Biochem. 266:892-902(1999).
DR EMBL; AF115745; AAD13109.1;
DR EMBL; AF109401; AAC98690.1;

1998
Dec.
1020

medline
1999
Dec
1020

Db 188 SPPCRPTRYEAVSFMDVNSTWRTVD 213

```

RESULT 5
O88839          PRELIMINARY;          PRT; 815 AA.
ID O88839
AC O88839
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE METALLOPROTEASE-DISINTEGRIN MDC15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98421554; PubMed=9748307;
RA Lum L., Reid M.S., Blobel C.P.;
RT "Intracellular maturation of the mouse metalloprotease disintegrin
RT MDC15."
RL J. Biol. Chem. 273:26236-26247(1998).
DR EMBL; AF006196; AAC61896.1; -.
DR HSP; P17494; 1KST.
DR MEROPS; M12.215; -.
DR InterPro; IPR000130; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR001590; -.
DR InterPro; IPR001762; -.
DR InterPro; IPR002870; -.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR ProDom; PD000664; -.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00050; DISIN; 1.
KW Integrin; Protease; Metalloprotease.
SQ SEQUENCE 815 AA; 87424 MW; C064BD3B7347D19B CRC64;

```

```

Query Match 7.2%; Score 10; DB 11; Length 815;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPA 13
   |||||
Db 800 PSRPAPPPA 809

```

```

RESULT 6
Q9QYV0          PRELIMINARY;          PRT; 816 AA.
ID Q9QYV0
AC Q9QYV0
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE METARGIDIN.
GN MDC15.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Bosse F., Petzold G., Greiner-Petter R., Pippirs U., Gillen C.,
RA Mueller H.;
RT "Cellular localization of the rat CRII-7/rMDC15 disintegrin mRNA in
RT PNS and CNS and regulated expression after nerve injury."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ251198; CAB61762.1; -.

```

```

DR HSP; P17494; 1KST.
DR InterPro; IPR000130; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR001590; -.
DR InterPro; IPR001762; -.
DR InterPro; IPR002870; -.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00050; DISIN; 1.
SQ SEQUENCE 816 AA; 88051 MW; B9D2CE023266FC27 CRC64;

```

```

Query Match 7.2%; Score 10; DB 11; Length 816;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 PSRPAPPPA 13
   |||||
Db 801 PSRPAPPPA 810

```

```

RESULT 7
O80640          PRELIMINARY;          PRT; 113 AA.
ID O80640
AC O80640
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE F12L6.19 PROTEIN.
GN F12L6.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC F12L6 genomic sequence."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004218; AAC27844.1; -.
SQ SEQUENCE 113 AA; 12641 MW; F9CB388D6AC083B4 CRC64;

```

```

Query Match 6.5%; Score 9; DB 10; Length 113;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 8 APPPPAPPS 16
   |||||
Db 2 APPPPAPPS 10

```

```

RESULT 8
Q9WNX1          PRELIMINARY;          PRT; 136 AA.
ID Q9WNX1
AC Q9WNX1
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE GAG/TAX FUSION PROTEIN (FRAGMENT).
OS Human T-cell lymphotropic virus type 1.
OC Viruses; Retroid viruses; Retroviridae; BLV-HTLV retroviruses.
OX NCBI_TaxID=11908;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Hill S.A., Shuh M., Dorse D.;
RT "Comparison of Defective HTLV-I Proviruses Predict the Mode of Origin
RT and Coding Potential of Internally Deleted Genomes.";
RL Virology 0:0-0(1999).
DR EMBL; AF148889; AAD44270.1; -.
DR InterPro; IPR000336; -.
DR InterPro; IPR002965; -.
DR InterPro; IPR003139; -.
DR Pfam; PF02228; gag_p19; 1.
DR PRINTS; PRO1217; PRICHEXTNSN.
DR PRODOM; PD001556; -. 1.
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 15240 MW; 4A8E6D241F3503F2 CRC64;

Query Match 6.5%; Score 9; DB 14; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
DB 77 PSRPAPPPP 85

RESULT 9
ID O05284 PRELIMINARY; PRT; 137 AA.
AC O05284;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 13.9 KDA PROTEIN.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RX MEDLINE=97169152; PubMed=9016959;
RA Donnenberg M.S., Lai L.C., Taylor K.A.;
RT "The locus of enterocyte effacement pathogenicity island of
RT enteropathogenic Escherichia coli encodes secretion functions and
RT remnants of transposons at its extreme right end.";
RL Gene 184:107-114(1997).
DR EMBL; L76581; AAB51428.1; -.
KW Hypothetical protein.
SQ SEQUENCE 137 AA; 13950 MW; 35C80CE8C6F403DC CRC64;

Query Match 6.5%; Score 9; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
DB 51 PSRPAPPPP 59

RESULT 10
ID Q82444 PRELIMINARY; PRT; 202 AA.
AC Q82444;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-WAR-2001 (TREMBLrel. 16, Last annotation update)
DE GAG POLYPROTEIN [CONTAINS: CORE PROTEIN(S) P24].
GN GAG OR P29GAG.
OS Human T-cell lymphotropic virus type 1.
OC Viruses; Retroid viruses; Retroviridae; BLV-HTLV retroviruses.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=H582;
RX MEDLINE=93376506; PubMed=8367298;
RA Orita S., Kobayashi H., Aono Y., Saiga A., Maeda M., Igarashi H.;
RT "p21x mRNA is expressed as a singly spliced px transcript from
RT defective provirus genomes having a partial deletion of the pol-env
RT region in human T-cell leukemia virus type 1-infected cells.";
RL Nucleic Acids Res. 2:3799-3807(1993).
CC -1- SIMILARITY: TO GAG GENE PROTEIN P24 (CORE NUCLEOCAPSID PROTEIN).
DR EMBL; D14335; BAA03265.1; -.
DR InterPro; IPR000721; -.
DR InterPro; IPR003139; -.
DR Pfam; PF00607; gag_p24; 1.
DR Pfam; PF02228; gag_p19; 1.
KW Core protein; Polyprotein.
SQ SEQUENCE 202 AA; 22381 MW; 00DA34CF0FB09F8 CRC64;

Query Match 6.5%; Score 9; DB 14; Length 202;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
DB 95 PSRPAPPPP 103

RESULT 11
ID O52152 PRELIMINARY; PRT; 206 AA.
AC O52152;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TREMBLrel. 06, Last annotation update)
DE ESPF.
GN ESPF.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Elliott S.J., Wainwright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K.,
RA Lai L.C., McNamara B.P., Donnenberg M.S., Kaper J.B.;
RL Mol. Microbiol. 0:0-0(1998).
DR EMBL; AF022236; AAC38400.1; -.
SQ SEQUENCE 206 AA; 20977 MW; 977341FB7D067950 CRC64;

Query Match 6.5%; Score 9; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
DB 73 PSRPAPPPP 81

RESULT 12
ID Q9RQEL PRELIMINARY; PRT; 207 AA.
AC Q9RQEL;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ESPF.
GN ESPF.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-B10;
RA Oswald E.;
RT "Right end of the locus of enterocyte effacement (LEE) from Rabbit
RT enteropathogenic Escherichia coli O103:H2.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF116900; AAF03351.1; -. 9586FB871E556726 CRC64;
SQ SEQUENCE 207 AA; 21192 MW; 9586FB871E556726 CRC64;

Query Match 6.5%; Score 9; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
DB 73 PSRPAPPPP 81
|||||

RESULT 13
O85622 PRELIMINARY; PRT; 248 AA.
AC O85622;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE L0016.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=EDL933;
RX MEDLINE=98339885; PubMed=9673266;
RA Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,
RA Kaper J.B., Blattner F.R.;
RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT Escherichia coli O157:H7.";
RL Infect. Immun. 66:3810-3817(1998).
DR EMBL; AF071034; AAC31495.1; -.
SQ SEQUENCE 248 AA; 25252 MW; 72949E5D24588B8B CRC64;

Query Match 6.5%; Score 9; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
DB 73 PSRPAPPPP 81
|||||

RESULT 14
Q9RD48 PRELIMINARY; PRT; 297 AA.
AC Q9RD48;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 31.2 KDA PROTEIN.
GN SCF36.19.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);

```

```

RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL MOL. Microbiol. 21:77-96(1996).
DR EMBL; ALI33424; CAB62764.1; -.
DR InterPro; IPR000379; -.
DR InterPro; IPR000734; -.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 297 AA; 31166 MW; 412FF60F1F9428D1 CRC64;

```

```

Query Match 6.5%; Score 9; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 32 SRARAAGAR 40
DB 215 SRARAAGAR 223
|||||

```

```

RESULT 15
O56229 PRELIMINARY; PRT; 428 AA.
ID O56229;
AC O56229;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE GAG POLYPROTEIN [CONTAINS: CORE PROTEIN(S) P24].
GN GAG.
OS Human T-cell lymphotropic virus type 1.
OC Viruses; Retroid viruses; Retroviridae; BLV-HTLV retroviruses.
OX NCBI_TaxID=11908;
[1]
RN SEQUENCE FROM N.A.
RA Petropoulos C.J.;
RL (In) Coffin J.M. (eds.);
RL Retroviruses, pp.757-757, Cold Spring Harbor Laboratory Press,
RL New York (1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Chappey C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE CCHC ZINC FINGER DOMAIN IS REQUIRED FOR VIRAL GENOME
CC PACKAGING AND FOR EARLY INFECTION PROCESS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
DR EMBL; AF033817; AAC82579.1; -.
DR InterPro; IPR000721; -.
DR InterPro; IPR001878; -.
DR InterPro; IPR001993; -.
DR InterPro; IPR003139; -.
DR Pfam; PF00098; zf-CCHC; 2.
DR Pfam; PF02228; gag_p24; 1.
DR Pfam; PF02228; gag_p19; 1.
DR PRINTS; PR00939; C2HCZFINGER.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR SMART; SM00343; znF_C2HC; 1.
KW Core protein; Polyprotein; Zinc-finger.
FT CHAIN 130 344 P24 CA.
FT CHAIN 345 428 P15NC-P14PR-P95RT/IN.
SQ SEQUENCE 428 AA; 47383 MW; 8B52787BC925BAAE CRC64;

Query Match 6.5%; Score 9; DB 14; Length 428;
Best Local Similarity 100.0%; Pred. No. 3;

```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 PSRPAPPPP 12
Db |||||
94 PSRPAPPPP 102

Search completed: August 16, 2001, 15:49:25
Job time: 285 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:38:30 ; Search time 21.79 Seconds
(without alignments)
395.031 Million cell updates/sec

Title: US-09-357-349-3

Perfect score: 601

Sequence: 1 AGPGSRARAGRCRLRS.....VNSTWRTVDRLSATACGCLG 113

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 232 | 38.6 | 197 | 2 T47159 | hypothetical prote |
| 2 | 174.5 | 29.0 | 211 | 2 T49686 | glial cell line-de |
| 3 | 169.5 | 28.2 | 211 | 2 A37499 | glial cell line-de |
| 4 | 167.5 | 27.9 | 211 | 2 B37499 | glial cell line-de |
| 5 | 108 | 18.0 | 560 | 1 WPHOM | mullerian inhibiti |
| 6 | 99 | 16.5 | 575 | 2 T11753 | mullerian inhibiti |
| 7 | 97 | 16.1 | 575 | 1 WFBOM | mullerian inhibiti |
| 8 | 91.5 | 15.2 | 553 | 1 A42499 | mullerian inhibiti |
| 9 | 88.5 | 14.7 | 555 | 1 S20100 | mullerian inhibiti |
| 10 | 87 | 14.5 | 350 | 2 T25451 | transforming growt |
| 11 | 86 | 14.3 | 238 | 2 T37040 | hypothetical prote |
| 12 | 84 | 14.0 | 151 | 2 S43296 | bone morphogenetic |
| 13 | 84 | 14.0 | 357 | 2 A39364 | GDF-1 embryonic gr |
| 14 | 83.5 | 13.9 | 309 | 2 JC5697 | placental transfor |
| 15 | 83.5 | 13.9 | 372 | 2 C39364 | GDF-1 embryonic gr |
| 16 | 82.5 | 13.7 | 616 | 2 QJ1441 | hypothetical 67K p |
| 17 | 81.5 | 13.6 | 641 | 1 QBE31 | nuclear antigen EB |
| 18 | 79 | 13.1 | 366 | 1 A42428 | inhibin alpha chai |
| 19 | 77.5 | 12.9 | 455 | 2 A43918 | TGF-beta-related p |
| 20 | 77.5 | 12.9 | 569 | 2 T50711 | urease (EC 3.5.1.5 |
| 21 | 75 | 12.5 | 255 | 2 S53099 | nep protein - huma |
| 22 | 74.5 | 12.4 | 216 | 2 T30657 | hypothetical prote |
| 23 | 73.5 | 12.2 | 407 | 2 T37242 | transforming growt |
| 24 | 73.5 | 12.2 | 409 | 2 S01825 | transforming growt |
| 25 | 72 | 12.0 | 255 | 1 A0LJ52 | nep protein - huma |
| 26 | 72 | 12.0 | 450 | 2 T01711 | probable serine/th |
| 27 | 72 | 12.0 | 1107 | 1 S52517 | myosin I heavy cha |
| 28 | 71.5 | 11.9 | 365 | 2 T43286 | cet-1 protein - Ca |
| 29 | 71.5 | 11.9 | 436 | 2 B54552 | cartilage-derived |

30 71 11.8 360 1 A25732 inhibin alpha chai
31 71 11.8 364 1 WPPGA inhibin alpha chai
32 71 11.8 364 2 F36470 Wnt-6 protein - mo
33 71 11.8 846 2 S52418 gtp-binding regula
34 71 11.8 879 2 I64133 phosphoenolpyruvat
35 70.5 11.7 410 2 A41397 transforming growt
36 70.5 11.7 728 2 T20561 transforming growt
37 70 11.6 411 1 I35604 platelet glycoprot
38 70 11.6 583 2 T16007 hypothetical prote
39 70 11.6 588 2 A26158 decapentaplegic pr
40 70 11.6 698 2 T17261 hypothetical prote
41 69.5 11.6 203 2 S32799 hypothetical prote
42 69.5 11.6 207 2 S37618 vgr protein - rat
43 69.5 11.6 312 2 A61183 hypothetical prote
44 69.5 11.6 410 2 A5706 transforming growt
45 69.5 11.6 412 2 A36169 transforming growt

ALIGNMENTS

RESULT 1

T47159

hypothetical protein DKFzp762B0211.1 - human

C:Species: Homo sapiens (man)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47159

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, March 2000

A:Reference number: 224379

A:Accession: T47159

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-197 <AAA>

A:Cross-references: EMBL:AL161995

A:Experimental source: adult melanoma (MeWo cell line); clone DKFzp762B0211

C:Genetics:

A:Note: DKFzp762B0211.1

Query Match 38.6%; Score 232; DB 2; Length 197;

Best Local Similarity 45.8%; Pred. No. 8.5e-17;

Matches 55; Conservative 10; Mismatches 41; Indels 14; Gaps 3;

Oy 4 PGSRARAA-----GARGCRLRSQLVPRVRAIGLCHRSDELVRFRFCGSCRRARSP 53

Db 81 PGPRRRGRPRRRRARRGARGPCGLRELVRSVSELGLGYASDEVLFVRYCAGACAAARV 140

Oy 54 HDLSIASLLGAGALRPPPGSRVSPQPCRPTRYE-AVSFMDVNSTWRTVDRLSATACGCL 112

Db 141 YDLGLRLRQRRLRLR---RERVRAQPCRPTRYEVSFLDAHSRYHTVHLSARECACV 197

RESULT 2

I49686

glial cell line-derived neurotrophic factor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 08-Oct-1999

C:Accession: I49686

R:Watabe, K.; Fukuda, T.; Tanaka, J.; Honda, H.; Toyohara, K.; Sakai, O.

J. Neurosci. Res. 41, 279-290, 1995

A:Title: Spontaneous immortalized adult mouse Schwann cells secrete autocrine and p

A:Reference number: I49686; MUID:95379105

A:Accession: I49686

A>Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: mRNA

A:Residues: 1-211 <RES>

A:Cross-references: GR:BA9972, NID:g758584; PIDN:BA08660.1; PID:g758585

R:Matsushita, N.; Fujita, Y.; Tanaka, M.; Nogatsu, T.; Kiuchi, K.

Gene 203, 149-157, 1997

A:Title: Cloning and structural organization of the gene encoding the mouse glial cel

A:Reference number: JC6518; MUID:98086214

A:Accession: JC6518

Superfamily:

| | | | | | | | | | |
|---------|-----|---------------------------------------|------|--------------------------------|-----|--------|-----|------|----|
| Matches | 35; | Conservative | 14; | Mismatches | 49; | Indels | 25; | Gaps | 7; |
| 2y | 3 | GPGPSARAAGARG-CRLRSQLVYPVALGLGHS | --- | DELVRPFCSGSC | --- | RRARSP | --- | 53 | |
| | | | | | | | | | |
| db | 441 | GRGRSKGTGTDGLCALRELSVDLRA | ---- | BRSLVLPETYQANNCQGACAWPQSDORNPY | | | | 496 | |
| 2y | 54 | --HDLSSLASLLGAGALRPPPGSRPVPQPCRPTRYEA | ---- | VSFMDVNSTWRTVDSLATA | | | | 108 | |

Db 497 GHHVLLKMQARGALG-----RLCCVPTATYTKLLISLSEEHISAHVENVVATE 549
 QY 109 CQC 111
 Db 550 CGC 552

RESULT 9
 S20100
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Dec-2000
 C:Accession: S20100; S51159
 R:Muensterberg, A.; Lovell-Badge, R.
 Development 113, 613-624, 1991
 A:Title: Expression of the mouse anti-Muellerian hormone gene suggests a role in both male and female sex differentiation
 A:Reference number: S20100; MUID:92146272
 A:Accession: S20100
 A:Molecule type: DNA
 A:Residues: 1-555 <MUE>
 A:Cross-references: EMBL:X63240; NID:g49945; PIDN:CAA44912.1; PID:g49946
 R:Dresser, D.W.; Hacker, A.; Lovell-Badge, R.; Guerrier, D.
 submitted to the EMBL Data Library, January 1995
 A:Description: The genes for anti-Muellerian hormone (AMH) and a spliceosome protein (SAH) are expressed in the developing testis.
 A:Reference number: S51159
 A:Accession: S51159
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-41 <DRE>
 A:Cross-references: EMBL:X83733
 C:Genetics:
 A:Introns: 135/1; 182/3; 219/1; 272/2
 C:Superfamily: inhibin

Query Match 14.7%; Score 88.5; DB 1; Length 555;
 Best Local Similarity 28.7%; Pred. No. 0.15;
 Matches 37; Conservative 10; Mismatches 51; Indels 31; Gaps 7;
 QY 3 GPGSRARAGARG-----CRLSQVPRVRLGLGHR--DELVRFRFCGSGCR--R 49
 Db 437 GREGRGRTAQRGDKGQDGFPCALRELSVDLRA-----ERSVLIPETVQANNCOGACRWPOS 492
 QY 50 ARSP-----HDLISLILGALGALRPPGSRPVSPCCRPTRYEA--VSPMDVNSTWRTVD 102
 Db 493 DRNPRTGNHVVLLKMQARGALG-----RLPCCVPTAYAGKLLISLSEERISADHVP 545
 QY 103 RLSATACGC 111
 Db 546 NMYATECGC 554

RESULT 10
 T25451
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
 C:Accession: T25451
 R:Bentley, D.
 submitted to the EMBL Data Library, December 1996
 A:Description: The sequence of C. elegans cosmid B0412.
 A:Reference number: Z20037
 A:Accession: T25451
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-350 <BEN>
 A:Cross-references: EMBL:U80953; PIDN:AAB52554.1; GSPDB:GN00021
 A:Experimental source: strain Bristol N2; clone B0412
 C:Genetics:
 A:Gene: daf-7
 A:Map position: 3
 A:Introns: 43/3; 123/3; 184/2; 288/3

C:Superfamily: inhibin

Query Match 14.5%; Score 87; DB 2; Length 350;
 Best Local Similarity 26.9%; Pred. No. 0.14;
 Matches 35; Conservative 16; Mismatches 47; Indels 32; Gaps 8;

QY 5 GSRAR-----AAGARGCRLRSQVPRVRLGLGHRSDLV---RFR--FCSGSRR 49
 Db 229 GSKRRSHAKPVCNAEAGSKGCLYDLETFEIKIGW----DWIVAPPRYNAYMCRGDCH- 283
 QY 50 ARSPHDLISLASLIGAGALRPPGSRPVSPQ-----CCRPTRYEAVSFMDVNSTWR----TV 101
 Db 284 -YNAHFNLAEATGHSKIMR---AAHKVSNPEIGYCCHPTEYDIKLIYVNRGRVSIANV 339
 QY 102 DRLSATACGC 111
 Db 340 NGMIATKCGC 349

RESULT 11
 T37040
 hypothetical protein SCJ12.23 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37040
 R:Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, August 1999
 A:Reference number: Z21619
 A:Accession: T37040
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-238 <MUR>
 A:Cross-references: EMBL:AL109989; PIDN:CAB53435.1; GSPDB:GN00070; SCOEDB:SCJ12.23
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SCJ12.23

Query Match 14.3%; Score 86; DB 2; Length 238;
 Best Local Similarity 35.1%; Pred. No. 0.12;
 Matches 39; Conservative 8; Mismatches 48; Indels 16; Gaps 6;
 QY 9 RAAGARGCRLRSQVPRVRLGLGHRSDLVRFRCGSGCR--ARSPHDLISLASLIGAGA 66
 Db 2 RGARARGORAGGHRLPVG---GRRPDHVRVRPGGAGSGARDGIARDVHDLVIQRLF-AGA 56
 QY 67 LRPPPG-SRPVSQPCCRPTRYEAVSFMD-----VNST---WRTVDRLSATA 108
 Db 57 LSPQALGRVTGRPRASERIORVVADLDDTIKIVIRSTIHALRESDRQTGTA 107

RESULT 12
 S43296
 bone morphogenetic protein-related protein (GDF7) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 19-May-2000
 C:Accession: S43296
 R:Storm, E.E.; Huynh, T.V.; Copeland, N.G.; Jenkins, N.A.; Kingsley, D.M.; Lee, S.J.
 Nature 368, 639-643, 1994
 A:Title: Limb alterations in brachypodism mice due to mutations in a new member of the bone morphogenetic protein family
 A:Reference number: S43294; MUID:94195427
 A:Accession: S43296
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-151 <STO>
 A:Cross-references: GB:U08339; NID:g488465; PIDN:AAA18780.1; PID:g488466
 C:Superfamily: inhibin

Query Match 14.0%; Score 84; DB 2; Length 151;
 Best Local Similarity 26.8%; Pred. No. 0.13;
 Matches 34; Conservative 16; Mismatches 51; Indels 26; Gaps 8;

| | | | | |
|-----------------------|--------|------------------|----------------|--------------------|
| Query Match | 13.9%; | Score 83.5; | DB 2; | Length 309; |
| Best Local Similarity | 28.8%; | Pred. No. 0.29; | | |
| Matches | 34; | Conservative 11; | Mismatches 50; | Indels 23; Gaps 6; |

Search completed: August 16, 2001, 15:40:05
Job time: 95 sec

C;Comment: This protein plays a role in reproduction.

Best Local Similarity 26.8%; PRed. NO: 0.29;
Matches 34: Conservative 11: Mismatches 50: Indels 22: Cons 6.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:40:47 ; Search time 35.14 Seconds
(without alignments)
523.347 Million cell updates/sec

Title: US-09-357-349-4
Perfect score: 747
Sequence: 1 PPOSRPAPPAPPALPR.....VNSTWTVDRLSATACGCLG 139

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTREMBL_16.*
- 2: sp_archaea.*
- 3: sp_bacteria.*
- 4: sp_fungi.*
- 5: sp_human.*
- 6: sp_invertebrate.*
- 7: sp_mammal.*
- 8: sp_mhc.*
- 9: sp_organelle.*
- 10: sp_phage.*
- 11: sp_plant.*
- 12: sp_rodent.*
- 13: sp_unclassified.*
- 14: sp_vertebrate.*
- 15: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 747 | 100.0 | 220 | 4 | Q96030 |
| 2 | 747 | 100.0 | 237 | 4 | Q95441 |
| 3 | 615 | 82.3 | 224 | 11 | Q920L2 |
| 4 | 459 | 61.4 | 125 | 11 | Q90ZG3 |
| 5 | 183 | 24.5 | 161 | 11 | Q90ZG0 |
| 6 | 182.5 | 24.4 | 185 | 11 | Q64063 |
| 7 | 182.5 | 24.4 | 211 | 11 | Q64062 |
| 8 | 181.5 | 24.3 | 160 | 6 | Q97685 |
| 9 | 178.5 | 23.9 | 133 | 4 | Q9UD32 |
| 10 | 175.5 | 23.5 | 185 | 4 | Q9UD33 |
| 11 | 152.5 | 20.4 | 182 | 13 | Q9IAM2 |
| 12 | 152.5 | 20.4 | 215 | 13 | Q9IAM3 |
| 13 | 107.5 | 14.4 | 512 | 10 | Q9LH25 |
| 14 | 105 | 14.1 | 364 | 13 | Q9PVK1 |
| 15 | 104.5 | 14.0 | 222 | 4 | Q9UF25 |
| 16 | 104.5 | 14.0 | 259 | 4 | Q9NP7 |
| 17 | 104.5 | 14.0 | 362 | 4 | Q95884 |
| 18 | 104.5 | 14.0 | 393 | 11 | O08719 |
| 19 | 104.5 | 14.0 | 418 | 4 | Q9UIC2 |

| | | | | | |
|----|-------|------|------|----|--------|
| 20 | 104.5 | 14.0 | 3247 | 14 | Q65553 |
| 21 | 104 | 13.9 | 521 | 6 | Q9TTC2 |
| 22 | 102.5 | 13.7 | 258 | 4 | Q9NPY5 |
| 23 | 101.5 | 13.6 | 393 | 11 | P70429 |
| 24 | 101.5 | 13.6 | 414 | 11 | Q9ERU8 |
| 25 | 101.5 | 13.6 | 1958 | 14 | Q69340 |
| 26 | 101 | 13.5 | 233 | 2 | Q9X8N0 |
| 27 | 101 | 13.5 | 261 | 11 | Q9EOT3 |
| 28 | 100 | 13.4 | 460 | 10 | Q9FWH4 |
| 29 | 99.5 | 13.3 | 36 | 11 | Q9JMC0 |
| 30 | 99.5 | 13.3 | 416 | 4 | Q9U108 |
| 31 | 99 | 13.3 | 571 | 4 | Q9UF76 |
| 32 | 99 | 13.3 | 609 | 4 | Q9NZN7 |
| 33 | 99 | 13.3 | 753 | 4 | Q75175 |
| 34 | 98 | 13.1 | 542 | 2 | O68872 |
| 35 | 98 | 13.1 | 658 | 10 | O65214 |
| 36 | 97.5 | 13.1 | 212 | 2 | O69990 |
| 37 | 97.5 | 13.1 | 480 | 11 | Q9QYU3 |
| 38 | 97 | 13.0 | 721 | 10 | Q9FF15 |
| 39 | 97 | 13.0 | 814 | 4 | Q9HGM7 |
| 40 | 97 | 13.0 | 846 | 11 | O63803 |
| 41 | 95.5 | 12.8 | 445 | 4 | O75267 |
| 42 | 95.5 | 12.8 | 528 | 4 | O60743 |
| 43 | 95.5 | 12.8 | 889 | 4 | Q9UL51 |
| 44 | 95.5 | 12.8 | 889 | 4 | Q9UBS2 |
| 45 | 95.5 | 12.8 | 1685 | 4 | Q9UEM8 |

ALIGNMENTS

| | |
|--------|--|
| RESULT | 1 |
| O96030 | |
| ID | O96030 |
| AC | O96030; PRELIMINARY; PRT; 220 AA. |
| DT | 01-MAY-1999 (TrEMBLrel. 10, Created) |
| DT | 01-MAY-1999 (TrEMBLrel. 10, Last sequence update) |
| DT | 01-MAR-2001 (TrEMBLrel. 16, Last annotation update) |
| DE | ARTEMIN. |
| GN | ARTN OR EVN. |
| OS | Homo sapiens (Human). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| OX | NCBI_TaxID=9606; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE-99098182; PubMed-9883723; |
| RA | Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H., |
| RA | Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr., |
| RA | Milbrandt J.; |
| RT | "Artemin, a novel member of the GDNF ligand family, supports |
| RT | peripheral and central neurons and signals through the GFRalpha3-RET |
| RT | receptor complex."; |
| RL | Neuron 21:1291-1302(1998). |
| RN | [2] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE=BRIN; |
| RA | Hansen C., Blom N., Johansen T.E.; |
| RT | "Neublastin a novel member of the GDNF ligand family."; |
| RL | Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases. |
| RN | [3] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE-20050601; PubMed-10583383; |
| RA | Masure S., Geerts H., Cik M., Hoefnagel E., Van Den Kleboom G., |
| RA | Tuytelaars A., Harris S., Lesage A.S., Leyse J.E., van der Helm L., |
| RA | Verhaesselt P., Yon J., Gordon R.D.; |
| RT | "Enovin, a member of the glial cell line-derived neurotrophic factor |
| RT | (GDNF) family with growth promoting activity on neuronal cells. |
| RT | existence and tissue-specific expression of different splice |
| RT | variants."; |
| RL | Eur. J. Biochem. 266:892-902(1999). |
| DR | EMBL; AF113765; AAC98690.1; |
| DR | EMBL; AF109401; AAC98690.1; |

```

DR EMBL; AF120274; AAD21075.1; -.
DR EMBL; AJ245628; CAB52396.1; -.
DR HSP; Q07731; IAGQ.
DR InterPro; IPR001839; -.
DR SMART; SM00204; TGFb; 1.
KW Signal.
SQ SEQUENCE 220 AA; 22906 MW; C47754B19AADCFBB CRC64;

Query Match 100.0%; Score 747; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.9e-60;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPQSRPAPPAPPALPRGGRARAGGPGSRARAAGCRLRSQLVVPVRLGLGHS 60
DB 82 PPQSRPAPPAPPAPPALPRGGRARAGGPGSRARAAGCRLRSQLVVPVRLGLGHS 141
QY 61 DELVRFRCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRTTRYEAVSFMDV 120
DB 142 DELVRFRCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRTTRYEAVSFMDV 201
QY 121 NSTWRTVDRLSATACGCLG 139
DB 202 NSTWRTVDRLSATACGCLG 220

RESULT 2
O95441 PRELIMINARY; PRT; 237 AA.
AC O95441;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE ARTEMIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99098192; PubMed=9883723;
RA Balon R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,
RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,
RA Milbrandt J.;
RT "Artemin, a novel member of the GDNF ligand family, supports
RT peripheral and central neurons and signals through the GFRalpha3-RET
RT receptor complex.";
RL Neuron 21:1291-1302(1998).
DR EMBL; AF109402; AAC98691.1; -.
DR HSP; Q07731; IAGQ.
DR MGD; MGI:1333791; Artn.
DR InterPro; IPR001839; -.
DR SMART; SM00204; TGFb; 1.
FT CHAIN 112 224 NEUROTROPHIC FACTOR ARTEMIN.
SQ SEQUENCE 224 AA; 23726 MW; 3328FB794581DF0B CRC64;

Query Match 82.3%; Score 615; DB 11; Length 224;
Best Local Similarity 82.5%; Pred. No. 1.7e-48;
Matches 118; Conservative 3; Mismatches 18; Indels 4; Gaps 1;

QY 1 PPQSRPAPPAPPAP-PSALPRGGRAAAGGPGSRARAAGCRLRSQLVVPVRLGL 56
DB 82 PPQSPQAPPAPPAPPALQSPALRGARAAAGRTSSARTTDARGCRLRSQLVVPVRLGL 141
QY 57 GHRSDLVRFRCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRTTRYEAVS 116
DB 142 GHSDDELIRFCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRTTRYEAVS 201
QY 117 FMDVNSTWRTVDRLSATACGCLG 139
DB 202 FMDVNSTWRTVDRLSATACGCLG 224

RESULT 4
O9Q2G3 PRELIMINARY; PRT; 125 AA.
AC O9Q2G3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ARTEMIN (FRAGMENT).
GN ARTN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=COCHLEA, SUBSTANTIA NTGRA;
RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;
RT "Expression of neurturin, artemin, persephin and their receptors GFRA-
RT 2 and GFRA-3 in the mature rat cochlea.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184919; AAF01241.1; -.
DR HSP; Q07731; IAGQ.
FT NON_TER 1 125
SQ SEQUENCE 125 AA; 12983 MW; 8EDE626E4B83231 CRC64;

```


Query Match 61.4%; Score 459; DB 11; Length 125;
 Best Local Similarity 80.2%; Pred. No. 1.4e-34;
 Matches 89; Conservative 3; Mismatches 15; Indels 4; Gaps 1;

QY 1 PPOPSRAPPAPPPAP-----PSALPRGGRARAGCGSRRARAGCRLRSQVLPVPRALGL 56
 DB 15 PPOSPQAPPFPFALQSPPAALRGARARAGTRSSRARATDARGCRLRSQVLPVPRALGL 74
 QY 57 GHSDDELVRFRFCGSCRRARSRPHDSLASLLGAGALRPPGSRPVSQPC 107
 DB 75 GHSDDELIRFRFCGSCRRARSRPHDSLASLLGAGALRPPGSRPVSQPC 125

RESULT 5
 Q90ZG0 PRELIMINARY; PRT; 161 AA.

ID Q90ZG0
 AC Q90ZG0
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE NEURTURIN (FRAGMENT).
 GN NTN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=SPRAGUE-DAWLEY; TISSUE=COCHLEA, SUBSTANTIA NIGRA;
 RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;
 RT "Expression of neurturin, artemin, persephin and their receptors GFRA-
 2 and GFRA-3 in the mature rat cochlea."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF184922; AAF01244.1; -;
 DR HSSP; Q07731; IAGO.
 DR InterPro; IPR001839; -;
 DR NON_TER 1
 FT NON_TER 161
 FT NON_TER 161
 SQ SEQUENCE 161 AA; 18337 MW; E7ACECAB302A93A7 CRC64;

Query Match 24.5%; Score 183; DB 11; Length 161;
 Best Local Similarity 43.9%; Pred. No. 1.7e-09;
 Matches 47; Conservative 10; Mismatches 44; Indels 6; Gaps 3;

QY 13 APPSALPRGGRARAGCGSRRARAGCRLRSQVLPVPRALGLHRSDELVRFRFCGSG 72
 DB 60 SPWARPSPGR--RRAGPRRRRARRPGSRPCGLRELEVVRVSELGLGYTSDTVLFRYCAGA 117
 QY 73 CRRARSRPHDSLASLLGAGALRPPGSRPVSQPCRPTRYE-AVGF 118
 DB 118 CEAAIRYDGLRLRQRRVRK---ERVRAHPCCRPTRYAEVSEVL 161

RESULT 6
 Q64063 PRELIMINARY; PRT; 185 AA.

ID Q64063
 AC Q64063; Q63214;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-2001 (TREMBLrel. 16, Last annotation update)
 DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR GDNF555 (GDNF PROTEIN) (FRAGMENT).
 GN GDNF.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP MEDLINE=95203379; PubMed=7895811;

RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
 RA Bergman L.W.;
 RT "cDNA sequence and differential mRNA regulation of two forms of glial
 cell line-derived neurotrophic factor in Schwann cells and rat
 skeletal muscle."
 RL Exp. Neurol. 131:47-52(1995).
 RN [2]
 RP SEQUENCE OF 1-50 FROM N.A.
 RC STRAIN=WISTAR; TISSUE=KIDNEY;
 RX MEDLINE=95210610; PubMed=7696586;
 RA Suter-Crazzolara C., Unsicker K.;
 RT "GDNF is expressed in two forms in many tissues outside the CNS."
 RL NeuroReport 5:2486-2488(1994).
 RN [3]
 RP SEQUENCE OF 1-50 FROM N.A.
 RC STRAIN=WISTAR; TISSUE=KIDNEY;
 RA Suter-Crazzolara C.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; S75585; AAB33892.1; -;
 DR EMBL; X92495; CAA63237.1; -;
 DR HSSP; Q07731; IAGO.
 DR InterPro; IPR001839; -;
 DR SMART; SM00204; TGFB; 1.
 DR NON_TER 185
 FT NON_TER 185
 SQ SEQUENCE 185 AA; 20678 MW; DBBA7B5653D6C035 CRC64;

Query Match 24.4%; Score 182.5; DB 11; Length 185;
 Best Local Similarity 34.3%; Pred. No. 2.2e-09;
 Matches 47; Conservative 22; Mismatches 55; Indels 13; Gaps 3;

QY 11 PPAPPSALPRGGRARAG-----GPGSRRARAGCRLRSQVLPVPRALGLHRSDE 62
 DB 53 PDQKAAALPRKRNRQAAASPEKRGKRGKNGCVLTALHNVLTGLGYETKE 112
 QY 63 LVRFRCGSCRRARSRPHDSLASLLGAGALRPPGSRPVSQPCRPTRYE-EAVSEMDVN 121
 DB 113 ELIFRYCSCGCEAAETMYDKILKNLSRRLT----SDKVGQACCRPVAFDDLSFLDLS 168
 QY 122 STWRTVDRLSATACGCL 138
 DB 169 LVYHILKHSKRCGCI 185

RESULT 7
 Q64062 PRELIMINARY; PRT; 211 AA.

ID Q64062
 AC Q64062;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR GDNF633 (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP MEDLINE=95203379; PubMed=7895811;
 RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
 RA Bergman L.W.;
 RT "cDNA sequence and differential mRNA regulation of two forms of glial
 cell line-derived neurotrophic factor in Schwann cells and rat
 skeletal muscle."
 RL Exp. Neurol. 131:47-52(1995).
 DR EMBL; S75583; AAB33891.1; -;
 DR HSSP; Q07731; IAGO.
 DR InterPro; IPR001839; -;
 DR SMART; SM00204; TGFB; 1.
 DR NON_TER 211
 FT NON_TER 211
 SQ SEQUENCE 211 AA; 23549 MW; AF45E64869D955E CRC64;

| | | | | |
|--------|--|---|------|---------|
| QY | 122 | STWRTVDRLSATAGCL | 138 | |
| | : | : | | |
| | : | : | | |
| Db | 144 | LVIHILRKHSKRGCI | 160 | |
| RESULT | 9 | | | |
| Q9UD32 | | | | |
| ID | Q9UD32 | PRELIMINARY; | PRT; | 133 AA. |
| AC | Q9UD32; | | | |
| DT | 01-MAY-2000 | (TrEMBLrel. 13, Created) | | |
| DT | 01-MAY-2000 | (TrEMBLrel. 13, Last sequence update) | | |
| DT | 01-MAR-2001 | (TrEMBLrel. 16, Last annotation update) | | |
| DE | ASTROCYTE-DERIVED TROPIC FACTOR 2, ATF-2. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |

| | Query Match | 23.5% | Score 175.5; | DB 4; | Length 185; |
|----|-----------------------|---|--------------------|-------|--------------------------------|
| | Best Local Similarity | 34.3%; | Pred. No. 9.3e-09; | | |
| | Matches | 47; | Conservative | 20; | Mismatches 57; Indels 13; Gaps |
| Qy | 11 | PPAPPSALPRGGRAARAG-----CPGSRRAAGARGCRLSQLVPRALGGLGRHSD | 62 | | |
| | | : : : : : : | | | |
| Db | 53 | PDKOMAVLPFRERNRQAAANPENSRCGKRGRRGKNGCVLTAIHLNVTDLGLGTETKE | 112 | | |
| Qy | 63 | LVRPFCSGCRRAARSPHDLASLLCAGALRPPGSRPVSPQCCRPTRY-EAVSFMDVN | 121 | | |
| | | : : : : : : | | | |
| Db | 113 | ELIFRYCSGCDAAETTYDKILNLSNRRL----VSDKVGQACCRPIAFDDDLDFLDDN | 168 | | |
| Qy | 122 | STWFTVDRLSATAGCL | 138 | | |
| | | : : : | | | |
| Db | 169 | LVYHILRKHSAAKRCGI | 185 | | |

DR EMBL; AF082178; AAD52011.1; -.
DR HSP; P18075; IBMP.
DR InterPro; IPR001111; -.
DR InterPro; IPR001839; -.
DR InterPro; IPR002405; -.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGF-beta; 1.
DR PRINTS; PR00669; INHIBIN.
DR PRODOM; PD000357; -; 1.
DR PROSITE; PS00250; TGF-beta; 1.
DR SMART; SM00204; TGF-beta; 1.
KW Glycoprotein.
SQ SEQUENCE 364 AA; 40071 MW; 76565716FBEB78D8 CRC64;

Query Match 14.1%; Score 105; DB 13; Length 364;
Best Local Similarity 25.2%; Pred. No. 0.041;
Matches 36; Conservative 20; Mismatches 57; Indels 30; Gaps 5;
QY 20 RCGRAARAGGCGSR-----ARAAGARGCRLRSQLVVPVRLGLGHRSDLEL----- 63
DB 226 RGASLPMAQVPAQPRDPFPAKLSPRSRLDRLPQCORHPLSVDFEIGWSGHIISPRG 285
QY 64 VRFRCSCSC-----RRARSPhdLSLASLLGAGALRPPGSRPVSPCCRRTRYEAVSFM 118
DB 286 YNAYHRCGSCPPPLGEMRNPTNHATVQSIINALKL-----SEGVSPPCCVDPKLSINLL 340
QY 119 -----DVNSTWRTVDRLSATAGCC 137
DB 341 YFDDENVVLKQYDDMVAGSCGC 363

RESULT 15
Q9UF25 PRELIMINARY; PRT; 222 AA.
AC Q9UF25;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE HYPOTHETICAL 22.4 KDA PROTEIN (FRAGMENT).
GN DKFZP586G1721.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133642; CAB63763.1; -.
DR InterPro; IPR002965; -.
DR PRINTS; PR01217; PRICHEXTENS.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 222 AA; 22369 MW; D0AF3E9C1C95EF95 CRC64;

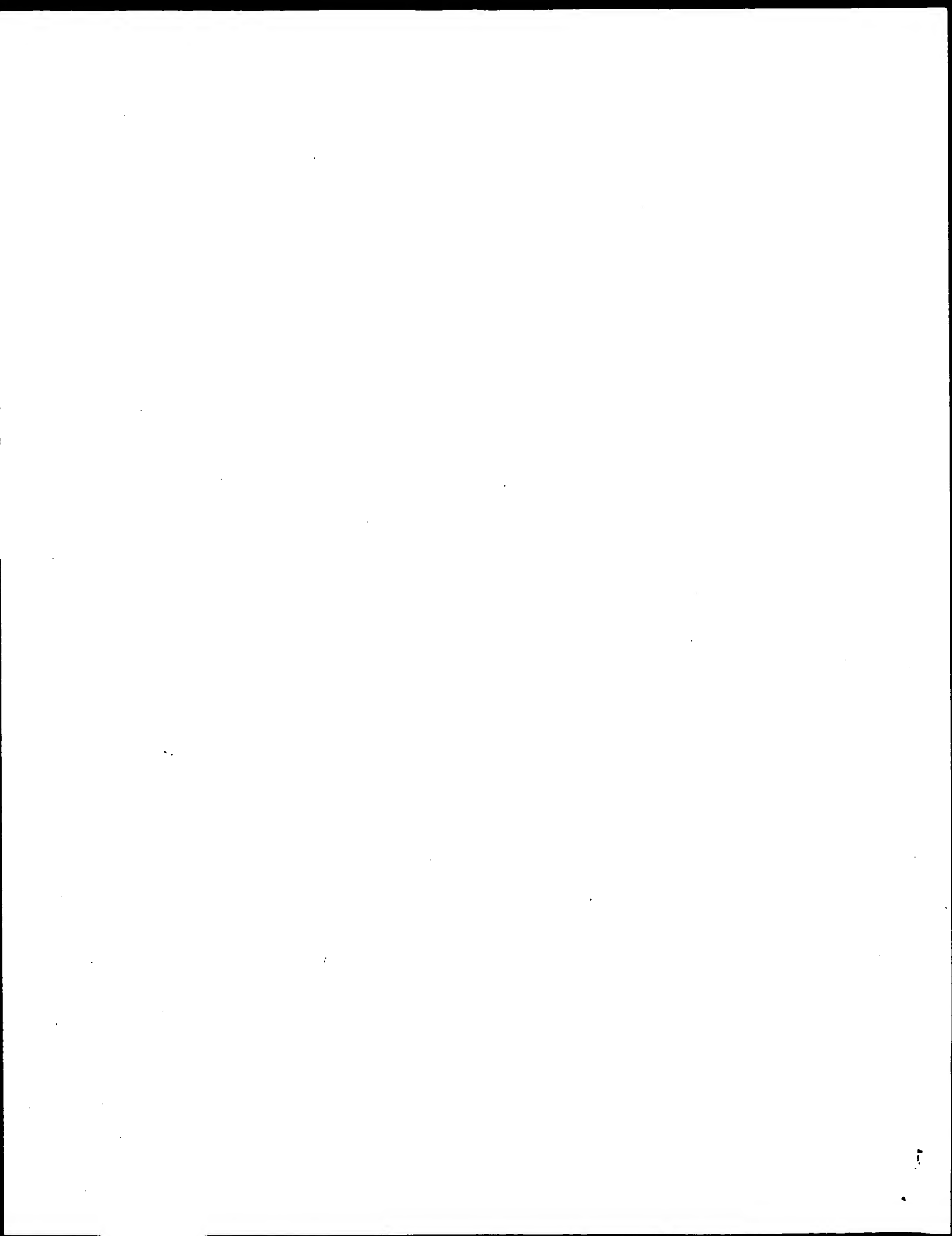
Query Match 14.0%; Score 104.5; DB 4; Length 222;
Best Local Similarity 26.1%; Pred. No. 0.029;
Matches 41; Conservative 6; Mismatches 49; Indels 61; Gaps 5;
QY 1 PPQSRPAPPPA-----PPSALPRGARAAGGPGSRA-RAAGARGCRLRSQLVVPVRL 54
DB 21 PPPPPPPPTGATPPPPPLPAGAGGSHDESSMSGLAAATAGAKLRVQRPEDAS 80
QY 55 GLGHRSDLVRRFCGSRARRSPHDLASLLGAG----- 91
DB 81 G-----GSSPSCSTSDANRASGGGGGLMEEMNKILAKRRKAASQSK 125
QY 92 -----ALRPPGSRPVSPCCRRTRYEA 114
DB 126 PAEKKEDESMEDPSTSPGTRAASQP---PNSSEA 159

Search completed: August 16, 2001, 15:40:48
Job time: 138 sec

• Fri Aug 17 09:56:42 2001

us-09-357-349-4.rspt

page 7



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:38:30 ; Search time 18.91 seconds
(without alignments)
123.041 Million cell updates/sec

Title: US-09-357-349-3
Perfect score: 601
Sequence: 1 AGGPGSRAAGACRCRLRS.....VNSTWRTVDRLSATACGCLG 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents:AA:*
1: /cgn2_6/ptodata/2/laa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/laa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/laa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/laa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/laa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/laa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 237 | 39.4 | 133 | 4 | US-08-931-858E-132 |
| 2 | 237 | 39.4 | 156 | 4 | US-08-931-858E-217 |
| 3 | 232 | 38.6 | 144 | 3 | US-08-775-414-81 |
| 4 | 232 | 38.6 | 152 | 3 | US-08-775-414-83 |
| 5 | 232 | 38.6 | 197 | 1 | US-08-519-777-7 |
| 6 | 232 | 38.6 | 197 | 1 | US-08-742-035-7 |
| 7 | 232 | 38.6 | 197 | 2 | US-08-777-019-7 |
| 8 | 232 | 38.6 | 197 | 2 | US-08-777-143-7 |
| 9 | 232 | 38.6 | 197 | 3 | US-09-106-486-5 |
| 10 | 232 | 38.6 | 197 | 3 | US-08-775-414-7 |
| 11 | 232 | 38.6 | 197 | 4 | US-08-931-858E-7 |
| 12 | 232 | 38.6 | 197 | 4 | US-08-981-739-7 |
| 13 | 228.5 | 38.0 | 134 | 4 | US-08-981-739-81 |
| 14 | 228.5 | 38.0 | 142 | 4 | US-08-931-858E-111 |
| 15 | 228.5 | 38.0 | 142 | 4 | US-08-981-739-111 |
| 16 | 228.5 | 38.0 | 156 | 4 | US-08-931-858E-185 |
| 17 | 228.5 | 38.0 | 185 | 4 | US-08-981-739-133 |
| 18 | 228 | 37.9 | 96 | 4 | US-08-931-858E-141 |
| 19 | 228 | 37.9 | 96 | 4 | US-08-981-739-141 |
| 20 | 227.5 | 37.9 | 96 | 4 | US-08-931-858E-80 |
| 21 | 227.5 | 37.9 | 96 | 4 | US-08-931-858E-187 |
| 22 | 227.5 | 37.9 | 96 | 4 | US-08-981-739-80 |
| 23 | 225.5 | 37.5 | 91 | 4 | US-08-931-858E-89 |
| 24 | 225.5 | 37.5 | 91 | 4 | US-08-981-739-89 |
| 25 | 225.5 | 37.5 | 96 | 4 | US-08-931-858E-221 |
| 26 | 225.5 | 37.5 | 104 | 3 | US-08-775-414-79 |
| 27 | 225 | 37.4 | 142 | 3 | US-08-775-414-82 |

| | | | | | | |
|----|-------|------|-----|---|--------------------|-------------------|
| 28 | 225 | 37.4 | 150 | 3 | US-08-775-414-84 | Sequence 84, Appl |
| 29 | 225 | 37.4 | 195 | 1 | US-08-519-777-8 | Sequence 8, Appl |
| 30 | 225 | 37.4 | 195 | 1 | US-08-742-035-8 | Sequence 8, Appl |
| 31 | 225 | 37.4 | 195 | 2 | US-08-777-019-8 | Sequence 8, Appl |
| 32 | 225 | 37.4 | 195 | 2 | US-08-777-143-8 | Sequence 8, Appl |
| 33 | 225 | 37.4 | 195 | 3 | US-08-775-414-8 | Sequence 8, Appl |
| 34 | 225 | 37.4 | 195 | 4 | US-08-931-858E-8 | Sequence 8, Appl |
| 35 | 225 | 37.4 | 195 | 4 | US-08-981-739-8 | Sequence 8, Appl |
| 36 | 223.5 | 37.2 | 89 | 4 | US-08-931-858E-79 | Sequence 79, Appl |
| 37 | 223.5 | 37.2 | 89 | 4 | US-08-931-858E-223 | Sequence 223, App |
| 38 | 223.5 | 37.2 | 89 | 4 | US-08-981-739-79 | Sequence 79, Appl |
| 39 | 222 | 36.9 | 102 | 1 | US-08-519-777-1 | Sequence 1, Appl |
| 40 | 222 | 36.9 | 102 | 1 | US-08-742-035-1 | Sequence 1, Appl |
| 41 | 222 | 36.9 | 102 | 2 | US-08-777-019-1 | Sequence 1, Appl |
| 42 | 222 | 36.9 | 102 | 2 | US-08-777-143-1 | Sequence 1, Appl |
| 43 | 222 | 36.9 | 102 | 3 | US-09-106-486-1 | Sequence 1, Appl |
| 44 | 222 | 36.9 | 102 | 3 | US-08-775-414-1 | Sequence 1, Appl |
| 45 | 222 | 36.9 | 102 | 4 | US-08-931-858E-1 | Sequence 1, Appl |

ALIGNMENTS

RESULT 1
US-08-931-858E-132
Sequence 132, Application US/08931858E
Patent No. 6222022
GENERAL INFORMATION:
APPLICANT: JOHNSON, EUGENE M
APPLICANT: MILBRANDT, JEFFREY D
APPLICANT: KOTZBAUER, PAUL T
APPLICANT: LAMPE, PATRICIA A
APPLICANT: KLEIN, ROBERT
APPLICANT: DESAUVAGE, FRED
TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
NUMBER OF SEQUENCES: 739
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931.858E
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 971486
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 132:
SEQUENCE CHARACTERISTICS:
LENGTH: 133 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-931-858E-132

Query Match 39.4% Score 237; DB 4; Length 133;
Best Local Similarity 45.5%; Pred. No. 7.9e-21;
Matches 55; Conservative 15; Mismatches 35; Indels 16; Gaps 3;

[illegible]

```

US-08-777-143-7
? Sequence 7, Application US/08777143
? Patent No. 5843914
? GENERAL INFORMATION:
? APPLICANT: JOHNSON JR., EUGENE M.
? APPLICANT: MILBRANDT, JEFFREY D.
? APPLICANT: KOTZBAUER, PAUL T.
? APPLICANT: LAMPE, PATRICIA A.
? TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
? NUMBER OF SEQUENCES: 78
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.
? STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
? CITY: ST. LOUIS
? STATE: MISSOURI
? COUNTRY: US
? ZIP: 63105-1817
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/777,143
? FILING DATE: 30-DEC-1996
? CLASSIFICATION: 514
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/519,777
? FILING DATE: 28-AUG-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: HOLLAND, DONALD R.
? REGISTRATION NUMBER: 35,197
? REFERENCE/DOCKET NUMBER: 953095
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (314) 727-5188
? TELEFAX: (314) 727-6032
? INFORMATION FOR SEQ ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 197 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-777-143-7

Query Match      38.6%; Score 232; DB 2; Length 197;
Best Local Similarity 45.8%; Pred. No. 5e-20;
Matches .55; Conservative 10; Mismatches 41; Indels 14; Gaps

Qy   4   PGSRARA-----GARGRLRSOLVPVRALGLGHRSDLVRFRFCSGSCRRARSP   53
      ||| |||||               ||||| ||||| ||||| ||||| ||||| |
Db   81  PGPRRRAGP RRNRARRLGRAPCGLRELVRSVELGLGVASDTVLFRYCAGACAAARV   140
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy   54  HDLSLASLLGAGALRP PPGSRPV SQPCRTRYE -AVSFMDVNSTWRTVDRLSATACGG   112
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Gb   141 YDLGLRLRQRRLR ----RRRVRAQQPCRTAYEDEVSFLDAHSRYHTVHELARECAV   197
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

```

US-09-106-486-5
Sequence 5, Application US/09106486
Patent No. 6043221
GENERAL INFORMATION:
APPLICANT: Magal, Ella
APPLICANT: Delaney, John M.
TITLE OF INVENTION: METHOD FOR PREVENTING AND TREATING
TITLE OF INVENTION: HEARING LOSS USING A NEURTURIN PROTEIN PRODUCT
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Angen Inc.
STREET: One Angen Center Drive
CITY: Thousand Oaks
STATE: California

RESULT 10
 US-08-778-414-7
 ; Sequence 7, Application US/08775414
 ; Patent No. 6090778
 ; GENERAL INFORMATION:
 ; APPLICANT: JOHNSON JR., EUGENE M.
 ; APPLICANT: MILBRANDT, JEFFREY D.
 ; APPLICANT: KOPFBAUER, PAUL T.
 ; APPLICANT: LAMPE, PATRICIA A.
 ; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
 ; NUMBER OF SEQUENCES: 90
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
 ; STREET: 7733 FORTSYTH BOULEVARD, SUITE 1400
 ; CITY: ST. LOUIS
 ; STATE: MISSOURI
 ; COUNTRY: US
 ; ZIP: 63105-1817
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/775,414
 ; FILING DATE: 31-DEC-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: HOLLAND, DONALD R.
 ; REGISTRATION NUMBER: 35,197
 ; REFERENCE/DOCKET NUMBER: 965805
 ; TELECOMMUNICATION INFORMATION:

Query Match 38.6%; Score 232; DB 4; Length 197;
Best Local Similarity 45.8%; Pred. No. 5e-20;
Matches 55; Conservative 10; Mismatches 41; Indels 14; Gaps 3;

QY 4 PGSRRAR-----GARGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSP 53
Db 81 PGPRRRAGRRRRARLARGPCGGLVRLVSEGLGYSADTVLFRYACAGCAAAARV 140
QY 54 HDLSLASLLGAGALRPPGSRPVSCCPRTRYE-AVSFMDVNSTWRTVDRLSATACGL 112
Db 141 YDLGLRLRQRRLR---RERVRAQPCCRPTAYEDVSFLDAHSRYHTVHLSARECAV 197

RESULT 12

US-08-981-739-7
; Sequence 7, Application US/08981739
; Patent No. 6232449
; GENERAL INFORMATION:
; APPLICANT: JOHNSON JR., EUGENE M.
; MILBRANDT, JEFFREY D.
; KOTZBAUER, PAUL T.
; LAMPE, PATRICIA A.
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: US
; ZIP: 63105-1817
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,739
; FILING DATE: 31-Aug-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/03461
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 976163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-08-981-739-7

Query Match 38.6%; Score 232; DB 4; Length 197;
Best Local Similarity 45.8%; Pred. No. 5e-20;
Matches 55; Conservative 10; Mismatches 41; Indels 14; Gaps 3;

QY 4 PGSRRAR-----GARGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSP 53
Db 81 PGPRRRAGRRRRARLARGPCGGLVRLVSEGLGYSADTVLFRYACAGCAAAARV 140
QY 54 HDLSLASLLGAGALRPPGSRPVSCCPRTRYE-AVSFMDVNSTWRTVDRLSATACGL 112
Db 141 YDLGLRLRQRRLR---RERVRAQPCCRPTAYEDVSFLDAHSRYHTVHLSARECAV 197

RESULT 13

US-08-981-739-81

; Sequence 81, Application US/08981739
; Patent No. 6232449
; GENERAL INFORMATION:
; APPLICANT: JOHNSON JR., EUGENE M.
; MILBRANDT, JEFFREY D.
; KOTZBAUER, PAUL T.
; LAMPE, PATRICIA A.
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: US
; ZIP: 63105-1817
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,739
; FILING DATE: 31-Aug-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/03461
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 976163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 81:

US-08-981-739-81

Query Match 38.0%; Score 228.5; DB 4; Length 134;
Best Local Similarity 46.3%; Pred. No. 8.1e-20;
Matches 50; Conservative 14; Mismatches 35; Indels 9; Gaps 2;

QY 7 RARAAGARGCRLRSQVPRALGLGHRSDLVRFRCGSC-RRARSPHDLASLLGAG 65
Db 35 RLPRALAGSCLRLSLTLPVAELGLGYASEKVIFFRYCAGSCPOEARTQHSVLRLRGRG 94
QY 66 ALRPPGSRPVSCCPRTRYE-AVSFMDVNSTWRTVDRLSATACGLG 113
Db 95 -----RAHGRPCQPTSYADVTFDLDQHHQOQLPQLSAAACGGG 134

RESULT 14

US-08-931-858E-111
; Sequence 111, Application US/08931858E
; Patent No. 6222022
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, EUGENE M
; MILBRANDT, JEFFREY D
; APPLICANT: KOTZBAUER, PAUL T
; APPLICANT: LAMPE, PATRICIA A
; APPLICANT: KLEIN, ROBERT
; APPLICANT: DESAUVAGE, FRED
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:

ADDRESSEE: HOWELL & HAERKAMP, L.C.
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,858E
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 971486
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
STRANDEDNESS: linear
MOLECULE TYPE: protein
US-08-931-858E-111

Query Match 38.0%; Score 228.5; DB 4; Length 142;
Best Local Similarity 46.3%; Pred. No. 8.7e-20;
Matches 50; Conservative 14; Mismatches 35; Indels 9; Gaps 2;
QY 7 RARAAGARCLRSQVPRALGLGHRSDLVRFRCGSC--RRARSPHDLASLLGAG 65
Db 43 RLPRALAGSCLWSLTLPLVAELGLGYASEEKVIFRYCAGSCPQEARQHSILVLRRLRG 102
QY 66 ALRPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
Db 103 -----RAHGRPCQPTSYADVTFLDDQHHWQQLPQLSAACGCGG 142

RESULT 15
US-08-981-739-111
Sequence 111, Application US/08981739
Patent No. 6232449
GENERAL INFORMATION:
APPLICANT: JOHNSON JR., EUGENE M.
MILBRANDT, JEFFREY D.
KOTZBAUER, PAUL T.
LAMPE, PATRICIA A.
TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
NUMBER OF SEQUENCES: 176
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAERKAMP, L.C.
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: US
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/981,739
FILING DATE: 31-Aug-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/03461
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 976163
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 111:
US-08-981-739-111

Query Match 38.0%; Score 228.5; DB 4; Length 142;
Best Local Similarity 46.3%; Pred. No. 8.7e-20;
Matches 50; Conservative 14; Mismatches 35; Indels 9; Gaps 2;
QY 7 RARAAGARCLRSQVPRALGLGHRSDLVRFRCGSC--RRARSPHDLASLLGAG 65
Db 43 RLPRALAGSCLWSLTLPLVAELGLGYASEEKVIFRYCAGSCPQEARQHSILVLRRLRG 102
QY 66 ALRPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
Db 103 -----RAHGRPCQPTSYADVTFLDDQHHWQQLPQLSAACGCGG 142

Search completed: August 16, 2001, 15:39:37
Job time: 67 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2001, 09:54:20 ; Search time 15.24 Seconds
(without alignments)
258.331 Million cell updates/sec

Title: US-09-357-349-3

Perfect score: 601

Sequence: 1 AGPGSPARAAGARGCRLRS.....VNSTWRTVDRLSATACGLG 113

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 94743 seqs, 34840360 residues

Total number of hits satisfying chosen parameters: 94743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 237 | 39.4 | 156 | 1 | PSPN_HUMAN |
| 2 | 232 | 38.6 | 197 | 1 | NRTN_HUMAN |
| 3 | 228.5 | 38.0 | 156 | 1 | PSPN_MOUSE |
| 4 | 225 | 37.4 | 195 | 1 | NRTN_MOUSE |
| 5 | 221.5 | 36.9 | 156 | 1 | PSPN_RAT |
| 6 | 174.5 | 29.0 | 211 | 1 | GDNF_MOUSE |
| 7 | 169.5 | 28.2 | 211 | 1 | GDNF_RAT |
| 8 | 167.5 | 27.9 | 211 | 1 | GDNF_HUMAN |
| 9 | 108 | 18.0 | 560 | 1 | MTS_HUMAN |
| 10 | 99 | 16.5 | 575 | 1 | MTS_PIG |
| 11 | 97 | 16.1 | 575 | 1 | MTS_BOVIN |
| 12 | 91.5 | 15.2 | 553 | 1 | MTS_RAT |
| 13 | 88.5 | 14.7 | 303 | 1 | GDF1_MOUSE |
| 14 | 88.5 | 14.7 | 555 | 1 | MTS_MOUSE |
| 15 | 87 | 14.5 | 350 | 1 | DAF7_CAEEL |
| 16 | 84 | 14.0 | 151 | 1 | GDF7_MOUSE |
| 17 | 84 | 14.0 | 357 | 1 | GDF1_MOUSE |
| 18 | 83.5 | 13.9 | 303 | 1 | GDF1_MOUSE |
| 19 | 83.5 | 13.9 | 308 | 1 | GDF1_HUMAN |
| 20 | 83.5 | 13.9 | 372 | 1 | GDF1_HUMAN |
| 21 | 81.5 | 13.6 | 641 | 1 | EBN1_EBV |
| 22 | 81 | 13.5 | 180 | 1 | NEF_HV2NV |
| 23 | 80.5 | 13.4 | 361 | 1 | IHA_TRIVU |
| 24 | 79 | 13.1 | 366 | 1 | IHA_HUMAN |
| 25 | 77.5 | 12.9 | 455 | 1 | 60A_DROME |
| 26 | 76.5 | 12.7 | 436 | 1 | 60A_DROVI |
| 27 | 74.5 | 12.4 | 1324 | 1 | IRS2_HUMAN |
| 28 | 73.5 | 12.2 | 399 | 1 | EB8A_MOUSE |
| 29 | 73.5 | 12.2 | 409 | 1 | TGF3_PIG |
| 30 | 73 | 12.1 | 367 | 1 | IHA_HORSE |
| 31 | 72.5 | 12.1 | 393 | 1 | BM15_SHEEP |
| 32 | 72 | 12.0 | 255 | 1 | NEF_HV2ST |
| 33 | 71.5 | 11.9 | 436 | 1 | GDF5_BOVIN |

ALIGNMENTS

RESULT 1

ID PSPN_HUMAN

AC 060542;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE PERSEPHIN PRECURSOR (PSP)

GN PSPN

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98150950; PubMed=9491986;

RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,

Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,

Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,

Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,

Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,

Phillips H.S., Johnson E.M.;

"Persephin, a novel neurotrophic factor related to GDNF and

neurturin."

RT Neuron 20:245-253(1998).

PL -1- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC

CC DOPAMINERGIC AND MOTOR NEURONS.

CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF040962; AAC39640.1; -

DR HSSP; Q07731; IAGQ.

DR TIM; 602921; -

DR InterPro: IPR001839; TGF-beta.

DR SMART; SM00204; TGFb; 1.

DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.

KW Growth factor; Signal.

FT SIGNAL 1 21

FT CHAIN 22 156

FT DISULFID 66 124

FT DISULFID 93 152

FT DISULFID 97 154

FT DISULFID 123 123

FT SEQUENCE 156 AA; 16600 MW; 6547751653A7044A CRC64;

INTERCHAIN (BY SIMILARITY).

POTENTIAL.

PERSEPHIN.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

Query Match

39.4%; Score 237; DB 1; Length 156;

DR EMBL: AL161995; CAB82327.1; -
DR HSP: Q07731; IAGO.
DR MIM: 602018; -
DR MIM: 142623; -
DR InterPro: IPRO02400; GF_cysknot.
DR InterPro: IPRO01839; TGF-beta.
DR PRINTS: PR00438; GFCYSKNOT.
DR SMART: SM00204; TGFB; 1.
DR PROSITE: PS00250; TGF_BETA_1; FALSE_NEG.
DR Growth factor; Signal; Disease mutation; Hirschsprung disease.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 95
FT CHAIN 96 197 BY SIMILARITY.
FT FT CHAIN 96 197 NEURTURIN.
FT DISULFID 103 165 BY SIMILARITY.
FT DISULFID 130 194 BY SIMILARITY.
FT DISULFID 134 196 BY SIMILARITY.
FT DISULFID 164 184 INTERCHAIN (BY SIMILARITY).
FT VARIANT 96 96 A -> S (in HSCR: ASSOCIATED TO A RET
FT MUTATION; INCOMPLETE PENETRANCE).
FT /FTID=VAR_009498.
FT SEQUENCE 197 AA; 22405 MW; 91AFAC8C3F8971FD CRC64;
Query Match 38.6%; Score 232; DB 1; Length 197;
Best Local Similarity 45.8%; Pred. No. 3.9e-17;
Matches 55; Conservative 10; Mismatches 41; Indels 14; Gaps 3;
QY 4 PGRARRAA-----GARGCLRSQVPRALGLHRSDELVRFRFGSGSCRRARP 53
DB 81 PGPRRAGPPRRRARARLARGPCGLEVRSELGLGYASDETVLFRCAGACEAARV 140
QY 54 HDLSLASLLGAGALRPPPGSRPSQPCCRPTRYE-AVSFMDVNSTWRTVDRLSATAGCL 112
DB 141 YDLGLRLRQRRLR---REVRVAQCPCRPAYEDEVSFLDAHSRYTHVELSARECAV 197
RESULT 3
PSPN_MOUSE STANDARD; PRT; 156 AA.
ID ID PSPN_MOUSE STANDARD; PRT; 156 AA.
AC AC 070300;
DT DT 30-MAY-2000 (Rel. 39, Created)
DT DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DE PERSEPHIN PRECURSOR (PSP).
GN GN PSPN.
OS OS Mus musculus (Mouse).
OC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
RN RN [1]
RP RP SEQUENCE FROM N.A.
RC RC STRAIN=129/SVJ;
RX RX MEDLINE=98150950; PubMed=9491986;
RA RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,
RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,
RA Stumpe K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,
RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,
RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,
RA Phillips H.S., Johnson E.M.;
RT RT "Persephin", a novel neurotrophic factor related to GDNF and
RT neuritin";
RL RL Neuron 20:245-253(1998).
CC CC FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC
CC CC DOPAMINERGIC AND MOTOR NEURONS.
CC CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC CC -1- SUBCELLULAR LOCATION: SECRETED.
CC CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF040960; AAC40057.1; --
 CC HSSP; Q07731; IAGO.
 DR MGD; MGI:1201684; Pspn.
 DR InterPro: IPR001839; TGF-beta.
 DR SMART; SM00204; TGF β ; 1.
 DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
 KW Growth factor; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 156 PERSEPHIN.
 FT DISULFID 66 124 BY SIMILARITY.
 FT DISULFID 93 152 BY SIMILARITY.
 FT DISULFID 97 154 BY SIMILARITY.
 FT DISULFID 123 123 INTERCHAIN (BY SIMILARITY).
 SQ SEQUENCE 156 AA; 17030 MW; 7DC6DD98132E041B CRC64;

Query Match 38.0%; Score 228.5; DB 1; Length 156;
 Best Local Similarity 46.3%; Pred. No. 7.1e-17;
 Matches 50; Conservative 14; Mismatches 35; Indels 9; Gaps 2;

QY 7 RARAAAGCRLRSQVLPVRAALGLHRSDELVRFCSCGSC-RRARSPHDLASLLGAG 65
 DB 57 RLPRALAGSCLWSLTPVPAELGLGYASEKVFYFCAGSCPQEARTOHSLVRLRGRG 116
 QY 66 ALRPPPGSPVSPCCPRTRYEAIVSMVNSTWRTVDRLSATACGLG 113
 DB 117 -----RAHGRCPCQTSYADYFLDDQHHWQQLPQLSAAACGGG 156

RESULT 4
 ID NRTN_MOUSE STANDARD; PRT; 195 AA.
 AC P97463;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE NEURTURIN PRECURSOR.
 GN NRTN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 96-110; 127-135; 155-177 & 181-190.
 RX MEDLINE=97100947; PubMed=8945474;
 RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,
 RA Creedon D.J., Johnson E.M. Jr., Milbrandt J.;
 RT "Neurturin, a relative of glial-cell-line-derived neurotrophic factor.";
 RL Nature 384:467-470(1996).

CC -!- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.
 CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT
 CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS
 CC HAEMOPOIETIC CELLS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- TISSUE SPECIFICITY: WIDESPREAD DISTRIBUTION.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U78109; AAC52954.1; --
 CC HSSP; Q07731; IAGO.
 DR MGD; MGI:108417; Nrtn.

DR InterPro: IPR002400; GF_cysknot.
 DR InterPro: IPR001839; TGF-beta.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR SMART; SM00204; TGF β ; 1.
 DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
 KW Growth factor; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 95 BY SIMILARITY.
 FT CHAIN 96 195 NEURTURIN.
 FT DISULFID 101 163 BY SIMILARITY.
 FT DISULFID 128 192 BY SIMILARITY.
 FT DISULFID 132 194 BY SIMILARITY.
 FT DISULFID 162 162 INTERCHAIN (BY SIMILARITY).
 SQ SEQUENCE 195 AA; 22219 MW; ABE21BB35D417448 CRC64;

Query Match 37.4%; Score 225; DB 1; Length 195;
 Best Local Similarity 45.9%; Pred. No. 2e-16;
 Matches 51; Conservative 12; Mismatches 44; Indels 4; Gaps 2;

QY 3 GPGSRARAAGCRLRSQVLPVRAALGLHRSDELVRFCSCGSCRRARSPHDLASLL 62
 DB 88 GPRRRARPGARPCGLRELVRSSELGTYTSDTVLFRYCAGACEAAIRIYDLGLRLR 147
 QY 63 GAGALRPPPGSPVSPCCPRTRYE-AVSEFMDVNSTWRTVDRLSATACGCL 112
 DB 148 QRRVR---RERARHPCCRTAYEDEVSLFVHRSYHTLQELSAECACV 195

RESULT 5
 ID PSPN_RAT STANDARD; PRT; 156 AA.
 AC O70301;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PERSEPHIN PRECURSOR (PSP).
 GN PSPN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98150950; PubMed=9491986;
 RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,
 RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,
 RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,
 RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,
 RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,
 RA Phillips H.S., Johnson E.M.;
 RT "Persephin, a novel neurotrophic factor related to GDNF and neurturin.";
 RL Neuron 20:245-253(1998).
 RN [2]
 RP SEQUENCE OF 1-78 FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Pons;
 RX MEDLINE=98374044; PubMed=9710270;
 RA Jaszi J., Farkas L.M., Galtier D., Reuss B., Strelau J., Unsicker K.,
 RA Krieglstein K.;
 RT "GDNF-related factor persephin is widely distributed throughout the nervous system.";
 RL J. Neurosci. Res. 53:494-501(1998).
 CC -!- FUNCTION: EXHIBITS NEUROTROPIC ACTIVITY ON MESENCEPHALIC
 CC DOPAMINERGIC AND MOTOR NEURONS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

| | |
|----|---|
| CC | EMBL; AF040961; AAC40058.1; -; |
| DR | EMBL; AJ005169; CAA06410.1; -; |
| DR | HSSP: Q07731; IAGQ. |
| DR | InterPro: IPR001839; TGF-beta. |
| DR | SMART: SM00204; TGFb: 1. |
| DR | PROSITE; PS00250; TGF_BETA_1; FALSE_NEG. |
| DR | Growth factor; Signal. |
| FT | SIGNAL 1 21 |
| FT | CHAIN 22 156 |
| FT | CHAIN 22 156 |
| FT | DISULFID 66 124 |
| FT | DISULFID 93 152 |
| FT | DISULFID 97 154 |
| FT | DISULFID 123 123 |
| FT | SEQUENCE 156 AA: 17063 MW: 9631941CCCB9B00B0 CRC64; |
| SO | INTERCHAIN (BY SIMILARITY). |

```

Query Match          36.9%; Score 221.5; DB 1; Length 156;
Best Local Similarity 45.5%; Pred. No. 3.7e-16;
Matches 45; Conservative 16; Mismatches 29; Indels 9; Gaps 12;

QY      16  CRLRSOLVVRALGLGHRSDLDLFRFCSCGSC--RRARSPHDLSLASLLGAGALRPPPGSR 74
         ||| :|| |||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB       66  CRLWSLTLPVALBGLGYASEEKILFYRCACGCPQEVRTQHSVLVARLGCG-----R 117

QY      75  PYVSQCPCRTRAEVFSFMDVNSTWRTVDRLSATACGCGLG 113
         :|||:||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB     118  AHGRPCCOPTSYADVTFDLDHHHMQQLPOLSAACGCG 156

RESULT    6
GNDF_MOUSE GNDF_MOUSE STANDARD: PRT: 211 AA.
ID GNDF_MOUSE STANDARD: PRT: 211 AA.
AC P48540; P97919; O09058; P97920; P70446;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR PRECURSOR.
```

[illegible]

```

RESULT      7
GNDF_RAT   STANDARD;          PRT;    211 AA.
ID          Q07731; Q64062; Q63214;
AC          01-FEB-1995 (Rel. 31, Created)
DT          01-FEB-1995 (Rel. 31, Last sequence update)
DT          01-OCT-2000 (Rel. 40, Last annotation update)
DE          GLIAL CELL LINE-DERIVED NEUTROPHIC FACTOR PRECURSOR.
DE          Rattus norvegicus (Rat).
GN          Rattus norvegicus (Rat).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX          NCBI_TaxID=10116;
RN          [1]
RP          SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 78-102.
RX          MEDLINE=93262463; PubMed=8493557;
RA          Lin L.-F.H., Doherty D.H., Tille J.D., Bektesh S., Collins F.;
RT          "GDNF: a glial cell line-derived neurotrophic factor for midbrain
RT          dopaminergic neurons.";

```

Science 260:1130-1132(1993).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=95203379; PubMed=7895811;
 RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
 Bergman L.W.;
 RT "cDNA sequence and differential mRNA regulation of two forms of glial
 cell line-derived neurotrophic factor in Schwann cells and rat
 skeletal muscle.";
 RL Exp. Neurol. 131:47-52(1995).
 RN [3]
 RP SEQUENCE OF 1-50 FROM N.A., AND ALTERNATIVE SPLICING.
 RC STRAIN=WTSTAR; TISSUE=Kidney;
 RX MEDLINE=95210610; PubMed=7696586;
 RA Suter-Crazzolara C., Unsicker K.;
 RT "GDNF is expressed in two forms in many tissues outside the CNS.";
 RL NeuroReport 5:2486-2488(1994).
 RN [4]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=95172201; PubMed=7867768;
 RA Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,
 Ramakrishnan L., Dreyfus C.F., Black I.B.;
 RT "Multiple astrocyte transcripts encode nigral trophic factors in rat
 and human.";
 RL Exp. Neurol. 130:387-393(1994).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=97331316; PubMed=9187648;
 RA Eigenbrot C., Gerber C.;
 RT "X-ray structure of glial cell-derived neurotrophic factor at 1.9-A
 resolution and implications for receptor binding.";
 RL Nat. Struct. Biol. 4:435-438(1997).
 CC -1- FUNCTION: NEUROTROPHIC FACTOR THAT ENHANCES SURVIVAL AND
 MORPHOLOGICAL DIFFERENTIATION OF DOPAMINERGIC NEURONS AND
 INCREASES THEIR HIGH-AFFINITY DOPAMINE UPTAKE.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1/GDNF633 (SHOWN HERE) AND
 2/GDNF555; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L15305; AAA67909.1; -;
 DR EMBL; S75583; AAB33891.1; -;
 DR EMBL; S75585; AAB33892.1; -;
 DR EMBL; X92495; CAA63237.1; -;
 DR PIR; A37499; A37499;
 DR PDB; 1AGO; 05-JUN-97.
 DR InterPro; IPR001839; TGF-beta.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
 KW Growth factor; Glycoprotein; Signal; Alternative splicing;
 KW 3D-structure.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 77
 FT CHAIN 78 211 GLIAL CELL LINE-DERIVED NEUROTROPHIC
 FT FACTOR.
 FT
 FT DISULFID 118 179
 FT DISULFID 145 208
 FT DISULFID 149 210
 FT DISULFID 178 178
 FT CARBOHYD 126 126
 FT CARBOHYD 162 162
 FT VARSPLIC 25 51
 FT
 FT CONFLICT 77 77 R -> S (IN REF. 2).
 FT

FT CONFLICT 90 90 E -> K (IN REF. 2).
 FT CONFLICT 101 101 E -> D (IN AA SEQUENCE).
 SQ SEQUENCE 211 AA; 23619 MW; AE06C64682895A5 CRC64;
 Query Match 28.2%; Score 169.5; DB 1; Length 211;
 Best Local Similarity 36.0%; Pred. No. 1.2e-10;
 Matches 40; Conservative 20; Mismatches 46; Indels 5; Gaps 2;
 QY 3 GPGSRARAGARGCRLRSOLVPRALGLGHSRDELVRFCGSCRRARSPHDLASLL 62
 DB 105 GKRRGQGRKNGRCVLTATHLNVTDLGLGYETKEELIFRYCGSCAAETMYDKILKLS 164
 QY 63 GAGALRPPPGSRPVSPQCCRPTRY-EAVSFMDVNSTWRTVDRLSATAGCCL 112
 DB 165 RSRLT----SDKVGQACRPVAFDDLSFLDLSLVYHLRKHSAKRCGCI 211
 RESULT 8
 GDNF_HUMAN
 ID GDNF_HUMAN STANDARD; PRT; 211 AA.
 AC P39905; Q9UP97; Q9UD33;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR PRECURSOR (ASTROCYTE-
 DE DERIVED TROPHIC FACTOR 1) (ATP-1).
 GN GDNF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=93262463; PubMed=8493557;
 RA Lin L.-F.H., Doherty D.H., Lille J.D., Bektesh S., Collins F.;
 RT "GDNF: a glial cell line-derived neurotrophic factor for midbrain
 dopaminergic neurons.";
 RL Science 260:1130-1132(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=95172201; PubMed=7867768;
 RA Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,
 Ramakrishnan L., Dreyfus C.F., Black I.B.;
 RT "Multiple astrocyte transcripts encode nigral trophic factors in rat
 and human.";
 RL Exp. Neurol. 130:387-393(1994).
 RN [3]
 RP SEQUENCE OF 1-187 FROM N.A. (ISOFORM 1).
 RC TISSUE=Kidney;
 RX MEDLINE=99296655; PubMed=10366742;
 RA Baecker P.A., Lee W.H., Verity A.N., Eglen R.M., Johnson R.M.;
 RT "Characterization of a promoter for the human glial cell line-derived
 neurotrophic factor gene.";
 RL Brain Res. Mol. Brain Res. 69:209-222(1999).
 RN [4]
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RX MEDLINE=97141760; PubMed=8988018;
 RA Hanu M., Hui J., Young Y., Le J., Katta V., Lee R., Shimamoto G.,
 Rohde M.F.;
 RT "Glial cell line-derived neurotrophic factor: selective reduction of
 the intermolecular disulfide linkage and characterization of its
 disulfide structure.";
 RL Biochemistry 35:16799-16805(1996).
 RN [5]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=98023959; PubMed=9359036;
 RA Hofstra R.M.W., Oslinga J., Buys C.H.C.M.;
 RT "Mutations in Hirschsprung disease: when does a mutation contribute to
 the phenotype.";
 RL Eur. J. Hum. Genet. 5:180-185(1997).
 RN [6]
 RP VARIANT HSCR SER-154.

RT Molecular genetics of the persistent mullerian duct syndrome: a study of 19 families.*;
 RL Hum. Mol. Genet. 3:125-131(1994).
 CC -!- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
 CC DUCT ORIGIN.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -!- DISEASE: DEFECTS IN AMH ARE THE CAUSE OF PERSISTENT MUELLERIAN
 CC DUCT SYNDROME TYPE I (PMDS-1): A FORM OF MALE
 CC PSEUDOHERMAPHRODITISM CHARACTERIZED BY A FAILURE OF MUELLERIAN
 CC DUCT REGRESSION IN OTHERWISE NORMAL MALES.
 CC -!- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR
 CC BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF
 CC RECEPTOR IN VITRO.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: K03474; AAC98805.1; -.
 CC EMBL: AC005263; AAC25614.1; -.
 CC PIR: A01397; WPHOM.
 CC MIM: 600957; -.
 CC MIM: 261550; -.
 CC InterPro: IPR002400; GF_CysKnot.
 CC InterPro: IPR001839; TGF-beta.
 CC Pfam: PF00019; TGF-beta; 1.
 CC PRINTS: PR00438; GFCYSKNOT.
 CC ProDom: PD000357; -. 1.
 CC SMART: SM00204; TGF-beta; 1.
 CC PROSITE: PS00250; TGF_BETA_1; 1.
 CC Growth factor; Glycoprotein; Gonadal differentiation; Signal;
 CC Pseudohermaphroditism; Disease mutation; Polymorphism.
 CC SIGNAL 1 18
 CC PROPEP 19 25
 CC CHAIN 26 560
 CC DISULFID 462 526
 CC DISULFID 488 557
 CC DISULFID 492 559
 CC DISULFID 525 525
 CC CARBOHYD 64 64
 CC CARBOHYD 329 329
 CC VARIANT 12 12
 CC
 CC VARIANT 49 49
 CC
 CC VARIANT 70 70
 CC
 CC VARIANT 101 101
 CC
 CC VARIANT 123 123
 CC
 CC VARIANT 167 167
 CC
 CC VARIANT 185 185
 CC
 CC VARIANT 194 194
 CC
 CC VARIANT 325 325
 CC
 CC VARIANT 477 477
 CC
 CC VARIANT 515 515
 CC
 CC CONFLICT 515 515
 CC SEQUENCE 560 AA; 59192 MW; 3EFC2EE4FEC364C CRC64;

Query Match

18.0%; Score 108; DB 1; Length 560;

Best Local Similarity 31.0%; Pred. No. 0.00068;
 Matches 39; Conservative 11; Mismatches 48; Indels 28; Gaps 7;
 QY 3 GPGSRARAAGARG-----CRLRSQVLPVRLGLGHRHRS-----DELVRFRCGSGC---RRARS 52
 DB 445 GPGRAQRAGATAADPCALRELSVDLRA-----ERSVLIPETQANNCGVCGWGPQSDRN 500
 QY 53 P-----HDLSSLALGAGALRPPGSRPVSPQCRPTRVEA---VSFMDVNSTWRTYVDRLS 105
 DB 501 PRYGHVHVLLKMQARGAALARP-----PCCVPTATAGKLLISLSEERISAHHPVNMV 553
 QY 106 ATACGC 111
 DB 554 ATECGC 559

RESULT 10

MIS_PIG
 ID MIS_PIG STANDARD; PRT; 575 AA.
 AC P79295;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)
 DE (AMH) (MULLERIAN INHIBITING SUBSTANCE).
 GN AMH.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Daneau I., Silversides D.W.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
 CC DUCT ORIGIN (BY SIMILARITY).
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U0853; AAC25968.1; -.
 CC InterPro: IPR001839; TGF-beta.
 CC Pfam: PF00019; TGF-beta; 1.
 CC ProDom: PD000357; -. 1.
 CC SMART: SM00204; TGF-beta; 1.
 CC PROSITE: PS00250; TGF_BETA_1; 1.
 CC Growth factor; Glycoprotein; Gonadal differentiation; Signal.
 CC SIGNAL 1 16
 CC PROPEP 17 23
 CC CHAIN 24 575
 CC DISULFID 477 541
 CC DISULFID 503 572
 CC DISULFID 507 574
 CC DISULFID 540 540
 CC CARBOHYD 78 78
 CC CARBOHYD 343 343
 CC SEQUENCE 575 AA; 61504 MW; 69AF63654B390780 CRC64;

Query Match

16.5%; Score 99; DB 1; Length 575;

Best Local Similarity 29.4%; Pred. No. 0.0059;
 Matches 37; Conservative 12; Mismatches 49; Indels 28; Gaps 7;

QY 3 GPGSRARAAGARG-----GCLRSQVLPVRLGLGHRHRS-----DELVRFRCGSGC---RRARS 52

DB 460 GPARAQRAGTAVSNGPCALRELSVDLRA-----ERSVLIPETYQANNCQGTGWPQSDRN 515
 QY 53 P-----HDLASLLGAGALRPPPGSRPVSPQCRPRTRYEA---VSFMDVNSTWRTVDRLS 105
 DB 516 PRYGNHVLLKMQARGAALARP-----PCCVPTAYAGKLLISLSERISAHHPVNMV 568
 QY 106 ATACGC 111
 DB 569 ATECGC 574

RESULT 11
 MIS_BOVIN STANDARD; PRT; 575 AA.
 ID MIS_BOVIN AC P03972;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 23-OCT-1986 (Rel. 02, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)
 DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).
 GN AMH.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=86218082; PubMed=3754790;
 RA Cate R.L., Mattaliano R.J., Hession C., Tizard R., Farber N.M.,
 RA Cheung A., Nifia E.G., Frey A.Z., Gash D.J., Chow E.P., Fisher R.A.,
 RA Bertonis J.M., Torres G., Wallner B.P., Ramachandran K.L.,
 RA Ragin R.C., Mangano T.F., McLaughlin D.T., Donahoe P.K.;
 RT "Isolation of the bovine and human genes for Mullerian inhibiting
 substance and expression of the human gene in animal cells.";
 RL Cell 45:685-698(1986).
 CC -1- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
 CC DUCT ORIGIN.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR
 CC BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF
 CC RECEPTOR IN VITRO.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M13151; AAA98765.1;
 CC PIR; A01398; WFBOM.
 CC InterPro: IPR002400; GF_cysknot.
 CC InterPro: IPR001839; TGF-beta.
 CC Pfam: PF00019; TGF-beta; 1.
 CC PRINTS; PR00438; GFCYSKNOT.
 CC ProDom: PD000357; -; 1.
 CC SMART; SM00204; TGFb; 1.
 CC PROSITE; PS00250; TGF_BETA.1; 1.
 CC Growth factor; Glycoprotein; Gonadal differentiation; Signal.
 FT SIGNAL 1 17
 FT PROPEP 18 24
 FT CHAIN 25 575
 FT CHAIN 25 575
 FT DISULFID 477 541
 FT DISULFID 503 572
 FT DISULFID 507 574
 FT DISULFID 540 574
 FT CARBOHYD 78 78
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 575 AA; 60623 MW; 892B89C11AC8B5A8 CRC64;
 Query Match 16.1%; Score 97; DB 1; Length 575;
 Best Local Similarity 29.4%; Pred. No. 0.0095;
 Matches 37; Conservative 12; Mismatches 49; Indels 28; Gaps 7;

QY 3 GPGSRARAAGARG-----CRLRSOLVPVRALGLGHRHRS---DELVRFRFCGSGC---RRARS 52
 DB 460 GSARAQRSAGAAAADGPCALRELSVDLRA-----ERSVLIPETYQANNCQACGWPQSDRN 515
 QY 53 P-----HDLASLLGAGALRPPPGSRPVSPQCRPRTRYEA---VSFMDVNSTWRTVDRLS 105
 DB 516 PRYGNHVLLKMQARGAALARP-----PCCVPTAYTGKLLISLSERISAHHPVNMV 568
 QY 106 ATACGC 111
 DB 569 ATECGC 574

RESULT 12
 MIS_RAT STANDARD; PRT; 553 AA.
 ID MIS_RAT AC P49000;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)
 DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).
 GN AMH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92241861; PubMed=1572639;
 RA Haqq C., Lee M.M., Tizard R., Wysz M., Demarinis J., Donahoe P.K.,
 RA Cate R.L.;
 RT "Isolation of the rat gene for Mullerian inhibiting substance.";
 RL Genomics 12:665-669(1992).
 CC -1- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
 CC DUCT ORIGIN.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S98336; AAB22104.1;
 CC InterPro: IPR002400; GF_cysknot.
 CC InterPro: IPR001839; TGF-beta.
 CC Pfam: PF00019; TGF-beta; 1.
 CC PRINTS; PR00438; GFCYSKNOT.
 CC ProDom: PD000357; -; 1.
 CC SMART; SM00204; TGFb; 1.
 CC PROSITE; PS00250; TGF_BETA.1; 1.
 CC Growth factor; Glycoprotein; Gonadal differentiation; Signal.
 FT SIGNAL 1 21
 FT PROPEP 22 ?
 FT CHAIN ? 553
 FT CHAIN ? 553
 FT DISULFID 455 519
 FT DISULFID 481 550
 FT DISULFID 485 552
 FT DISULFID 518 518
 FT INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 325 325 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 553 AA; 5888 MW; 75DAF3949A038A69 CRC64;

Query Match 15.2%; Score 91.5; DB 1; Length 553;
 Best Local Similarity 28.5%; Pred. No. 0.034;
 Matches 35; Conservative 14; Mismatches 49; Indels 25; Gaps 7;

QY 3 GPGSRARAGARG-CRLRSOLVPRALGLHRS---DELVRFRFCGSC---RRARSP-- 53
 Db 441 GRAGRSKGTGDLGALRELSVDLRA---ERSVLIPETVQANNCOGACAWQSDRNPY 496

QY 54 --HDLASLLGAGALRPPGSRPVSPQCRPTRYEA---VSEMDVNSTWRTVDRLSATA 108
 Db 497 GNHVLLKMQARGALG-----RLPCCVPTAYTKLLISLSEHISAHVHPNNVATE 549

QY 109 CGC 111
 Db 550 CGC 552

RESULT 13
 GDFE_RAT STANDARD; PRT; 303 AA.
 ID Q92076; AC Q92076;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE GROWTH/DIFFERENTIATION FACTOR 15 PRECURSOR (GDF-15).
 GN GDF15 OR SBF.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boettner M., Laaff M., Suter-Crazzolara C.;
 RT Identification of a novel member of the TGFbeta superfamily. "
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ011969; CAA09891.1;
 DR EMBL; AJ011970; CAA09891.1; JOINED.
 DR HSSP; P18075; 1BMP.
 DR InterPro: IPR001839; TGF-beta.
 DR Pfam: PF00019; TGF-beta; 1.
 DR ProDom: PD000357; -; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF-BETA; FALSE NEG.
 KW Growth factor; Cytokine; Glycoprotein; Signal.
 FT SIGNAL 1 30
 FT PROPEP 31 188
 FT CHAIN 189 303
 FT DISULFID 206 269
 FT DISULFID 235 300
 FT DISULFID 239 302
 FT DISULFID 268 268
 FT CARBOHYD 71 71
 SQ SEQUENCE 303 AA; 33438 MW; FA828BE79C1052C9 CRC64;

Query Match 14.7%; Score 88.5; DB 1; Length 303;

Best Local Similarity 29.9%; Pred. No. 0.038;
 Matches 35; Conservative 10; Mismatches 49; Indels 23; Gaps 6;

QY 3 GPGSRARAGARG-CRLRSOLVPRALGLGH---RSELDVFRFCGSC---RRARSPHLS 57
 Db 201 GPG-----RCCHLETVQATLEDLGSWDWVLSRQLQSLMCGECPHLYRSANTHALI 252

QY 58 LASLLGAGALRPPGSRPVSPQCRPTRYEAVSPM---DVNSTWRTVDRLSATAGC 111
 Db 253 KARLHG---LQPD-----RVPAPCCVPSSYTPVLMHRTDSGVSLOTYDDLVAQCCHC 302

RESULT 14
 MIS_MOUSE STANDARD; PRT; 555 AA.
 ID P27106; AC P27106;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)
 DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).
 GN AMH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129; TISSUE=Testis;
 RX MEDLINE=92146272; PubMed=1782869;
 RA Munsterberg A., Lovell-Badge R.;
 RT "Expression of the mouse anti-mullerian hormone gene suggests a role
 RT in both male and female sexual differentiation.";
 RL Development 113:613-624(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Dresser D.W., Jamn S., Atkins C.J., Guerrier D.;
 RT "A GNRP-like gene shares a bidirectional promoter with SAP62
 RT immediately upstream of AMH.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-42 FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=96081226; PubMed=8541848;
 RA Dresser D.W., Hacker A., Lovell-Badge R., Guerrier D.;
 RT "The genes for a spliceosome protein (SAP62) and the anti-Mullerian
 RT hormone (AMH) are contiguous.";
 RL Hum. Mol. Genet. 4:1613-1618(1995).
 CC -1- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
 CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
 CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
 CC DUCT ORIGIN
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -1- TISSUE SPECIFICITY: SERTOLI CELLS OF FETAL TESTES, AND TESTES
 CC JUST AFTER BIRTH, BUT ABSENT IN ADULT TESTES. IN FEMALE, AMH
 CC IS EXPRESSED AFTER BIRTH IN THE GRANULOSA CELLS OF THE FOLLICLE.
 CC AMH EXPRESSION IS DEPENDENT ON THE DEGREE OF FOLLICULAR
 CC MATURATION AND NOT ON THE AGE OF THE OVARY.
 CC -1- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR
 CC BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF
 CC RECEPTOR IN VITRO.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X63240; CAA44912.1; -.

DR EMBL; X83733; CAC10450.1; -.
DR PIR; S20100; S20100.
DR MGD; MGI:88006; Amh.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGF-beta.
DR Pfam; PF00019; TGF-beta; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR ProDom; PD000357; -. 1.
DR SMART; SM00204; TGF-beta; 1.
DR PROSITE; PS00250; TGF-BETA_1; 1.
KW Growth factor; Glycoprotein; Gonadal differentiation; Signal.
FT SIGNAL 1 20
FT PROPEP 21 ?
FT CHAIN 21 555
FT DISULFID 457 521
FT DISULFID 483 552
FT DISULFID 487 554
FT DISULFID 520 520
FT CARBOHYD 62 62
FT CARBOHYD 326 326
FT CARBOHYD 410 410
SQ SEQUENCE 555 AA; 59778 MW; DGA3A20C50306E29 CRC64;

Query Match 14.7%; Score 88.5; DB 1; Length 555;
Best Local Similarity 28.7%; Pred. No. 0.069;
Matches 37; Conservative 10; Mismatches 51; Indels 31; Gaps 7;

QY 3 GPGSRARAAGARG-----CRLRSQVVRALGLGHR-----DELVRFRFGSGCR---R 49
DB 437 GREGGRTRAQRGQDGQPCALRELSVDLRA-----ERSVLIPETYQANNCGACRWPQS 492
QY 50 ARSP-----HDLISLALGLGALRPPGSRPVQPCCRTRYEA---VSFMDVNSTWRTVD 102
DB 493 DRNPVGNHVILLKMQARGAALG-----RLPCCVPTAYAGKLLISLSEBISADHVP 545
QY 103 RLSATACGC 111
DB 546 NWVATECGC 554

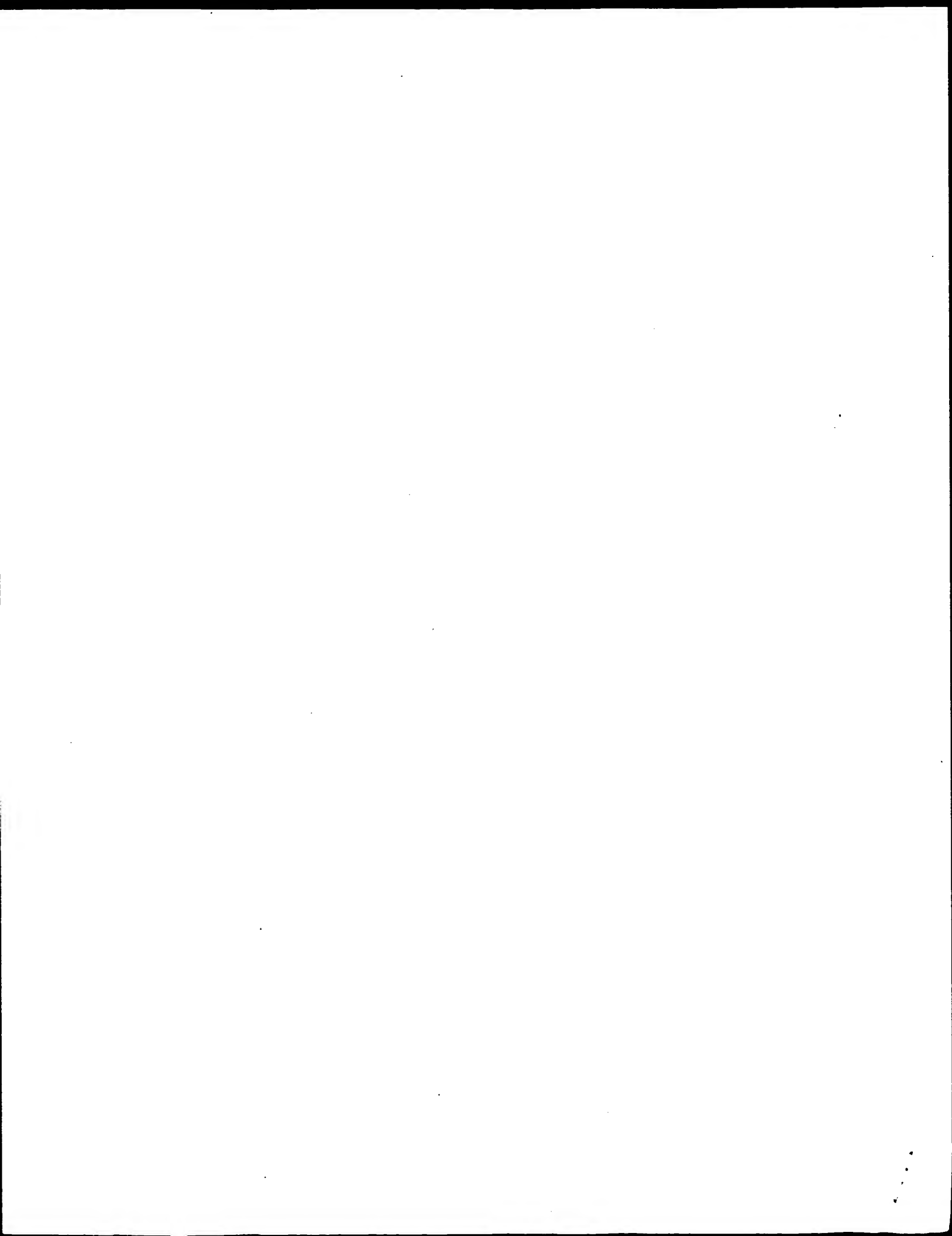
RESULT 15
DAF7_CABEL STANDARD; PRT; 350 AA.
AC P92172;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DAUER LARVA DEVELOPMENT REGULATORY GROWTH FACTOR DAF-7 PRECURSOR.
GN DAF-7 OR B0412.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=97067238; PubMed=8910282;
RA Ren P., Lim C.-S., Johnsen R., Albert P.S., pilgrim D., Riddle D.L.;
RT "Control of C. elegans larval development by neuronal expression of a
TGF-beta homolog.";
RL Science 274:1389-1391(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Bentley D.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBSJ databases.
CC -1- FUNCTION: MAY ACT AS A NEGATIVE REGULATOR OF DAUER LARVA
DEVELOPMENT BY TRANSDUCING CHEMOSENSORY INFORMATION FROM ASI
NEURONS.
CC -1- TISSUE SPECIFICITY: EXPRESSION IN THE CHEMOSENSORY NEURONS.
CC -1- DEVELOPMENTAL STAGE: WHEN THE FOOD/PEROMONE RATIO IS HIGH, ITS
LEVEL PEAKS DURING THE L1 LARVAL STAGE. EXPRESSION IS DETECTED IN

CC LARVAE BEGINNING 4 TO 5 HOURS AFTER HATCHING, THROUGH THE FOUR
CC LARVAL STAGES, AND IN ADULTS.
CC -1- INDUCTION: DAUER-INDUCING PHEROMONE INHIBITS ITS EXPRESSION AND
CC PROMOTES DAUER FORMATION, WHEREAS FOOD REACTIVATES ITS EXPRESSION
CC AND PROMOTES RECOVERY FROM THE DAUER STATE.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U72883; AAC47390.1; -.
CC EMBL; U72884; AAC47389.1; -.
CC EMBL; U80953; AAB52554.1; -.
CC HSSP; P18075; LBMP.
CC WormPep; B0412.2; CE01758.
CC InterPro; IPR001839; TGF-beta.
CC Pfam; PF00019; TGF-beta; 1.
CC ProDom; PD000357; -. 1.
CC SMART; SM00204; TGF-beta; 1.
CC PROSITE; PS00250; TGF-BETA_1; 1.
KW Growth factor; Developmental
FT SIGNAL 1 21
FT PROPEP 22 234
FT CHAIN 235 350
FT DISULFID 241 251
FT DISULFID 250 315
FT DISULFID 278 347
FT DISULFID 282 349
FT DISULFID 314 314
FT CARBOHYD 23 23
FT MUTAGEN 271 271
FT MUTAGEN 280 280
SQ SEQUENCE 350 AA; 39533 MW; F957C70A2B1FDE0A CRC64;

Query Match 14.5%; Score 87; DB 1; Length 350;
Best Local Similarity 26.9%; Pred. No. 0.063;
Matches 35; Conservative 16; Mismatches 47; Indels 32; Gaps 8;

QY 5 GSRAR-----AAGARGCRLRSQVVRALGLGHRSELY---RFR--FCSGSCRR 49
DB 229 GSKRRSHAKPVCNAEQSKGCLYDLIEFEKIGW----DWIVAPPRYNAYMCRGDCH- 283
QY 50 ARSPHDLISLALGLGALRPPGSRPVQPCCRTRYEA---CCRPTRYEAVSFMDVNSTWR---TV 101
DB 284 -YNAHHFNLAETGHSKIMR-----AAHKVSNPEIGYCCHPTDYIKLIYVNRDGRVSIANV 339
QY 102 RLSATACGC 111
DB 340 NGMIKKKGC 349

Search completed: August 17, 2001, 09:54:43
Job time: 23 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:38:30 ; Search time 35.14 Seconds
(without alignments)
425.454 Million cell updates/sec

Title: US-09-357-349-3
Perfect score: 601
Sequence: 1 AGGCGSRAARAAGRCRLRS.....VNSTWRTVDRLSATACGLG 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SP-REMBL16.*
- 2: sp-archaea.*
- 3: sp-bacteria.*
- 4: sp-fungi.*
- 5: sp-human.*
- 6: sp-invertebrate.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-unclassified.*
- 13: sp-vertebrate.*
- 14: sp-virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1 | 601 | 100.0 | 220 | 4 | O96030 | O96030 homo sapien |
| 2 | 601 | 100.0 | 237 | 4 | O95441 | O95441 homo sapien |
| 3 | 528 | 87.9 | 224 | 11 | Q9Z012 | Q9Z012 mus musculu |
| 4 | 372 | 61.9 | 125 | 11 | Q9Z0G3 | Q9Z0G3 rattus norv |
| 5 | 173.5 | 28.9 | 160 | 6 | O97685 | O97685 macaca mula |
| 6 | 171 | 28.5 | 161 | 11 | Q9Z0G0 | Q9Z0G0 rattus norv |
| 7 | 169.5 | 28.2 | 185 | 11 | O64063 | O64063 rattus norv |
| 8 | 169.5 | 28.2 | 211 | 11 | O64062 | O64062 rattus norv |
| 9 | 167.5 | 27.9 | 133 | 4 | Q9UD32 | Q9UD32 homo sapien |
| 10 | 167.5 | 27.9 | 185 | 4 | Q9UD33 | Q9UD33 homo sapien |
| 11 | 152.5 | 25.4 | 182 | 13 | Q9IAM2 | Q9IAM2 gallus gall |
| 12 | 152.5 | 25.4 | 215 | 13 | Q9IAM3 | Q9IAM3 gallus gall |
| 13 | 99.5 | 16.6 | 36 | 11 | O9JMC0 | O9JMC0 rattus norv |
| 14 | 98 | 16.3 | 364 | 13 | Q9PVK1 | Q9PVK1 gallus gall |
| 15 | 86 | 14.3 | 238 | 2 | Q9RI37 | Q9RI37 streptomyce |
| 16 | 85 | 14.1 | 512 | 10 | Q9LH25 | Q9LH25 oryza sativ |
| 17 | 84 | 14.0 | 390 | 13 | Q91597 | Q91597 xenopus lae |
| 18 | 83 | 13.8 | 257 | 4 | Q9H7T3 | Q9H7T3 homo sapien |
| 19 | 82.5 | 13.7 | 616 | 2 | Q49182 | Q49182 mycobacteri |

ALIGNMENTS

RESULT 1

| | | | |
|---|--------------|------|---------|
| O96030 | PRELIMINARY; | PRT; | 220 AA. |
| AC O96030: | | | |
| DT 01-MAY-1999 (TRENBLrel. 10, Created) | | | |
| DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update) | | | |
| DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update) | | | |
| DE ARTEMIN. | | | |
| GN ARTN OR EVN. | | | |
| OS Homo sapiens (Human). | | | |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| OX NCBI_TaxID=9606; | | | |
| RN [1] | | | |
| RP SEQUENCE FROM N.A. | | | |
| RX MEDLINE=95098192; PubMed=9883723; | | | |
| RA Balch R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H., | | | |
| RA Simburger K.S., Leltner M.L., Araki T., Johnson E.M. Jr., | | | |
| RA Milbrandt J.; | | | |
| RT "Artemin, a novel member of the GDNF ligand family, supports | | | |
| RT peripheral and central neurons and signals through the GFRalpha3-RET | | | |
| RT receptor complex."; | | | |
| RL Neuron 21:1291-1302(1998). | | | |
| RN [2] | | | |
| RP SEQUENCE FROM N.A. | | | |
| RA Hansen C., Blom N., Johansen T.E.; | | | |
| RA Tissue-BRAIN; | | | |
| RL "Neublastin a novel member of the GDNF ligand family."; | | | |
| RN Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RN [3] | | | |
| RP SEQUENCE FROM N.A. | | | |
| RX MEDLINE=20050601; PubMed=10583383; | | | |
| RA Measure S., Geerts H., Cik M., Hoefnagel E., Van Den Kieboom G., | | | |
| RA Tuytelaars A., Harris S., Lesage A.S., Leysen J.E., van der Helm L., | | | |
| RA Verhaesselt P., Von J., Gordon R.D.; | | | |
| RT "Enovin, a member of the glial cell-line-derived neurotrophic factor | | | |
| RT (GDNF) family with growth promoting activity on neuronal cells. | | | |
| RT Existence and tissue-specific expression of different splice | | | |
| RT variants."; | | | |
| RL Eur. J. Biochem. 266:892-902(1999). | | | |
| DR EMBL; AF115765; AAC13109.1; -. | | | |
| DR EMBL; AF109401; AAC98690.1; -. | | | |

Q9UJA2 homo sapien
O9YF2 rhodobacter
O9GLC6 rangifer ta
Q9QVR2 rattus sp.
Q34207 pseudomonas
P91720 drosophila
O46306 drosophila
Q76638 human immun
Q8XZ62 drosophila
Q9V4F4 drosophila
Q98223 molluscum c
Q35219 mus musculu
Q35221 mus musculu
O44371 caenorhabdi
Q9FP42 oryza sativ
Q9VQ99 drosophila
Q9YTT4 human immun
Q9GZW5 homo sapien
Q9V4E6 drosophila
O04601 arabidopsis
Q9SM80 oryza sativ
Q63356 rattus norv
Q9Y615 homo sapien
Q81344 hepatitis e
Q81862 hepatitis e
Q89444 hepatitis e

20 80 13.3 301 4 Q9UJA2
21 77.5 12.9 569 2 Q9RFF2
22 77 12.8 440 6 Q9GLC6
23 77 12.8 937 11 Q9QVR2
24 76.5 12.7 275 2 Q34207
25 76.5 12.7 614 5 P91720
26 75 12.5 139 5 O46306
27 75 12.5 255 14 Q76638
28 75 12.5 598 5 Q8XZ62
29 75 12.5 598 5 Q9V4F4
30 74.5 12.4 216 14 Q98223
31 74 12.3 1162 11 Q35219
32 74 12.3 1162 11 Q35221
33 73.5 12.2 407 5 O44371
34 73.5 12.2 542 10 Q9FP42
35 73.5 12.2 586 5 Q9VQ99
36 73 12.1 262 14 Q9YTT4
37 73 12.1 306 4 Q9GZW5
38 73 12.1 701 5 Q9V4E6
39 72 12.0 450 10 O04601
40 72 12.0 1095 10 Q9SM80
41 72 12.0 1107 11 Q63356
42 72 12.0 1338 4 Q9Y615
43 72 12.0 1693 14 Q81344
44 72 12.0 1693 14 Q81862
45 72 12.0 1693 14 Q89444

DR EMBL; AF120274; AAD21075.1; -;
 DR EMBL; AJ245628; CAB52396.1; -;
 DR HSSP; Q07731; IAGO.
 DR InterPro; IPR001839; -;
 DR SMART; SM00204; TGFb; 1.
 KW Signal.
 SQ SEQUENCE 220 AA; 22906 MW; C47754B19AADCFBB CRC64;

Query Match 100.0%; Score 601; DB 4; Length 220;
 Best Local Similarity 100.0%; Pred. No. 2.1e-55;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAARAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 60
 |||||
 DB 108 AGPGSRAARAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 167
 |||||

QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 |||||
 DB 168 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 220
 |||||

RESULT 2

ID O95441 PRELIMINARY; PRT; 237 AA.
 AC O95441;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-2001 (TrEMBLrel. 16, Last annotation update)
 DE ARTEMIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99098192; PubMed-9883723;
 RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,
 RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,
 RA Milbrandt J.;
 RT "Artemin, a novel member of the GDNF ligand family, supports
 peripheral and central neurons and signals through the GFRalpha3-RET
 receptor complex."
 RL Neuron 21:1291-1302(1998).
 DR EMBL; AF115765; AAD13110.1; -;
 DR HSSP; Q07731; IAGO.
 DR InterPro; IPR001839; -;
 DR SMART; SM00204; TGFb; 1.
 SQ SEQUENCE 237 AA; 24471 MW; 11C64C4B510CE3AB CRC64;

Query Match 100.0%; Score 601; DB 4; Length 237;
 Best Local Similarity 100.0%; Pred. No. 2.3e-55;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAARAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 60
 |||||
 DB 125 AGPGSRAARAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 184
 |||||

QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 |||||
 DB 185 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 237
 |||||

RESULT 3

ID Q920L2 PRELIMINARY; PRT; 224 AA.
 AC Q920L2;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE NEUROTROPHIC FACTOR ARTEMIN PRECURSOR.
 GN ARTN.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99098192; PubMed-9883723;
 RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,
 RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,
 RA Milbrandt J.;
 RT "Artemin, a novel member of the GDNF ligand family, supports
 peripheral and central neurons and signals through the GFRalpha3-RET
 receptor complex."
 RL Neuron 21:1291-1302(1998).
 DR EMBL; AF109402; AAC98691.1; -;
 DR HSSP; Q07731; IAGO.
 DR MGI; MGI:1333791; Artn.
 DR InterPro; IPR001839; -;
 DR SMART; SM00204; TGFb; 1.
 FT CHAIN 112 224 NEUROTROPHIC FACTOR ARTEMIN.
 SQ SEQUENCE 224 AA; 23726 MW; 3328FB794581DF0B CRC64;

Query Match 87.9%; Score 528; DB 11; Length 224;
 Best Local Similarity 88.5%; Pred. No. 9.4e-48;
 Matches 100; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGPGSRAARAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 60
 |||||
 DB 112 AGTRSSRARATDARGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 171
 |||||

QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 |||||
 DB 172 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 224
 |||||

RESULT 4

ID Q9QZG3 PRELIMINARY; PRT; 125 AA.
 AC Q9QZG3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ARTEMIN (FRAGMENT).
 GN ARTN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-COCHLEA, SUBSTANTIA NIGRA;
 RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;
 RT "Expression of neurturin, artemin, persephin and their receptors GFRA-
 2 and GFRA-3 in the mature rat cochlea."
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF184919; AAF01241.1; -;
 DR HSSP; Q07731; IAGO.
 FT NON_TER 1 1
 FT NON_TER 125 125
 SQ SEQUENCE 125 AA; 12983 MW; 8EDE626E4B83231 CRC64;

Query Match 61.9%; Score 372; DB 11; Length 125;
 Best Local Similarity 87.7%; Pred. No. 1.2e-31;
 Matches 71; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 AGPGSRAARAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 60
 |||||
 DB 45 AGTRSSRARATDARGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLSLAS 104
 |||||

QY 61 LLGAGALRPPPGSRPVSPQCC 81
 |||||

Db 105 LLGAGALRPPGSRPISQPC 125

RESULT 5

Q97685 PRELIMINARY; PRT; 160 AA.
 AC O97685;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE NEUROTROPHIC FACTOR (FRAGMENT).
 GN GDNF.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ErJia C., Tong L., QiuJiang D.;
 RT "The gene cloning of macaca and human GDNF by direct PCR from whole
 blood and sequence analysis."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF106678; AAC99782.1;
 DR HSSP; Q07731; IAGO.
 DR InterPro; IPR001839;
 DR SMART; SM00204; TGFb; 1.
 FT NON_TER 1
 SQ SEQUENCE 160 AA; 18196 MW; E206362185D499B4 CRC64;

Query Match 28.9%; Score 173.5; DB 6; Length 160;
 Best Local Similarity 37.8%; Pred. NO. 8.5e-11;
 Matches 42; Conservative 18; Mismatches 46; Indels 5; Gaps 2;

QY 3 GPGSRARAAGCRLRSQVPRALGLHRSDELVRFRCSCRRARSPHDLASLL 62
 Db 54 GKGRGCRGKNGCVLTALHNVTDLGLGYETKEELIFRYCSGDAAEITYDKILKNLS 113
 QY 63 GAGALRPPGSRPVQPCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112
 Db 114 RNRL-----VSDKVGACCRPTAFDDLSFLDNLVYLKHSKRCGCI 160

RESULT 6

Q90ZG0 PRELIMINARY; PRT; 161 AA.
 AC Q90ZG0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE NEURTURIN (FRAGMENT).
 GN NTN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=COCHLEA, SUBSTANTIA NIGRA;
 RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;
 RT "Expression of neurturin, artemin, persephin and their receptors GFRA-
 2 and GFRA-3 in the mature rat cochlea."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF184922; AAF01244.1;
 DR HSSP; Q07731; IAGO.
 DR InterPro; IPR001839;
 DR NON_TER 1
 FT NON_TER 161
 SQ SEQUENCE 161 AA; 18337 MW; E7ACECAB302A93A7 CRC64;

Query Match 28.5%; Score 171; DB 11; Length 161;

Best Local Similarity 45.1%; Pred. No. 1.6e-10;
 Matches 41; Conservative 9; Mismatches 37; Indels 4; Gaps 2;

QY 3 GPGSRARAAGCRLRSQVPRALGLHRSDELVRFRCSCRRARSPHDLASLL 62
 Db 74 GPRRRRARPGRPCGLRELVRSSELGYTSDETVLFYRCAGACAAIRYDGLRLRL 133
 QY 63 GAGALRPPGSRPVQPCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112
 Db 134 QRRVRK---ERVRAHPCCRPAYEDEVSF 161

RESULT 7

Q64063 PRELIMINARY; PRT; 185 AA.
 AC Q64063; Q63214;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR GDNF555 (GDNF PROTEIN)
 DE (FRAGMENT).
 GN GDNF.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95203379; PubMed=7895811;
 RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
 RA Bergman L.W.;
 RT "cDNA sequence and differential mRNA regulation of two forms of glial
 cell line-derived neurotrophic factor in Schwann cells and rat
 skeletal muscle."
 RT Exp. Neurol. 131:47-52(1995).
 RN [2]
 RP SEQUENCE OF 1-50 FROM N.A.
 RC STRAIN=WISTAR; TISSUE=KIDNEY;
 RX MEDLINE=95210610; PubMed=7696586;
 RA Suter-Crazzolara C., Unsicker K.;
 RT "GDNF is expressed in two forms in many tissues outside the CNS."
 RL NeuroReport 5:2486-2488(1994).
 RN [3]
 RP SEQUENCE OF 1-50 FROM N.A.
 RC STRAIN=WISTAR; TISSUE=KIDNEY;
 RA Suter-Crazzolara C.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; S75585; AAB33892.1;
 DR EMBL; X92495; CAA63237.1;
 DR HSSP; Q07731; IAGO.
 DR InterPro; IPR001839;
 DR SMART; SM00204; TGFb; 1.
 FT NON_TER 185
 SQ SEQUENCE 185 AA; 20678 MW; DBBA7B5653D6C035 CRC64;

Query Match 28.2%; Score 169.5; DB 11; Length 185;
 Best Local Similarity 36.0%; Pred. NO. 2.6e-10;
 Matches 40; Conservative 20; Mismatches 46; Indels 5; Gaps 2;

QY 3 GPGSRARAAGCRLRSQVPRALGLHRSDELVRFRCSCRRARSPHDLASLL 62
 Db 79 GKGRGCRGKNGCVLTALHNVTDLGLGYETKEELIFRYCSGDAAEITYDKILKNLS 138
 QY 63 GAGALRPPGSRPVQPCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112
 Db 139 RSRRLT----SDKVGACCRPVAFDDLSFLDNLVYLKHSKRCGCI 185

RESULT 8

Q64062 PRELIMINARY; PRT; 211 AA.
 ID Q64062
 AC Q64062;

DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR GDNF633 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95203379; PubMed=7895811;
RA Springer J.E., Seeburger J.L., He J., Gabrea A., Blankenhorn E.P.,
RA Bergman L.W.;
RT "CDNA sequence and differential mRNA regulation of two forms of glial
cell line-derived neurotrophic factor in Schwann cells and rat
skeletal muscle";
RL Exp. Neurol. 131:47-52(1995).
DR EMBL; S75583; AAB33891.1; -;
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -;
DR SMART; SM00204; TGFb; 1.
FT NON_TER 211 211
SQ SEQUENCE 211 AA; 23549 MW; AF4564869D955E CRC64;

Query Match 28.2%; Score 169.5; DB 11; Length 211;
Best Local Similarity 36.0%; Pred. No. 2.9e-10;
Matches 40; Conservative 20; Mismatches 46; Indels 5; Gaps 2;

QY 3 GPGSRAAGARGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLASLL 62
DB 105 GKRGQRCKRGKRGCVLTALHNLVDTGLGKETKEELIFRYCSGCEAAETVYDKILNLS 164

QY 63 GAGALRPPGSRPVSPQCCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112
DB 165 RSRLT-----SDKVGQACCRPVAFDDLSFLDNLVYHLRKHSAKRCGCI 211

RESULT 9
Q9UD32 PRELIMINARY; PRT; 133 AA.
AC Q9UD32;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ASTROCYTE-DERIVED TROPHIC FACTOR 2, ATF-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95172201; PubMed=7867768;
RA Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,
RA Ramakrishnan L., Dreyfus C.F., Black I.B.;
RT "Multiple astrocyte transcripts encode nigral trophic factors in rat
and human";
RL Exp. Neurol. 130:387-393(1994).
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -;
DR SMART; SM00204; TGFb; 1.
SQ SEQUENCE 133 AA; 14736 MW; B46B96DD5F679769 CRC64;

Query Match 27.9%; Score 167.5; DB 4; Length 133;
Best Local Similarity 36.9%; Pred. No. 3e-10;
Matches 41; Conservative 18; Mismatches 47; Indels 5; Gaps 2;

QY 3 GPGSRAAGARGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLASLL 62
DB 27 GKRGQRCKRGKRGCVLTALHNLVDTGLGKETKEELIFRYCSGCEAAETVYDKILNLS 86

QY 63 GAGALRPPGSRPVSPQCCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112

DB 87 RNRL-----VSDKVGQACCRPIAFDDLSFLDNLVYHLRKHSAKRCGCI 133

RESULT 10
Q9UD33 PRELIMINARY; PRT; 185 AA.
AC Q9UD33;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE ASTROCYTE-DERIVED TROPHIC FACTOR 1, ATF-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95172201; PubMed=7867768;
RA Schaar D.G., Sieber B.A., Sherwood A.C., Dean D., Mendoza G.,
RA Ramakrishnan L., Dreyfus C.F., Black I.B.;
RT "Multiple astrocyte transcripts encode nigral trophic factors in rat
and human";
RL Exp. Neurol. 130:387-393(1994).
DR HSSP; Q07731; IAGQ.
DR InterPro; IPR001839; -;
DR SMART; SM00204; TGFb; 1.
SQ SEQUENCE 185 AA; 20885 MW; 1988C50DA5EA1B10 CRC64;

Query Match 27.9%; Score 167.5; DB 4; Length 185;
Best Local Similarity 36.9%; Pred. No. 4.1e-10;
Matches 41; Conservative 18; Mismatches 47; Indels 5; Gaps 2;

QY 3 GPGSRAAGARGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLASLL 62
DB 79 GKRGQRCKRGKRGCVLTALHNLVDTGLGKETKEELIFRYCSGCEAAETVYDKILNLS 138

QY 63 GAGALRPPGSRPVSPQCCRPTRY-EAVSFMDVNSTWRTVDRLSATACGCL 112
DB 139 RNRL-----VSDKVGQACCRPIAFDDLSFLDNLVYHLRKHSAKRCGCI 185

RESULT 11
Q9IAM2 PRELIMINARY; PRT; 182 AA.
AC Q9IAM2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR SHORT FORM (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20092738; PubMed=10625540;
RA Homma S., Oppenheim R.W., Yaginuma H., Kimura S.;
RT "Expression pattern of GDNF, c-ret, and GFRalphas suggests novel roles
for GDNF ligands during early organogenesis in the chick embryo";
RL Dev. Biol. 217:121-137(2000).
DR EMBL; AF16018; AAF26685.1; -;
DR InterPro; IPR001839; -;
FT NON_TER 1 1
FT NON_TER 182 182
SQ SEQUENCE 182 AA; 20740 MW; 6A8AC16BD1B4F103 CRC64;

Query Match 25.4%; Score 152.5; DB 13; Length 182;
Best Local Similarity 35.2%; Pred. No. 1.5e-08;
Matches 38; Conservative 18; Mismatches 47; Indels 5; Gaps 2;

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE HYPOTHETICAL 24.9 KDA PROTEIN.
 GN SCJ12.23.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Murphy L., Harris D.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA MEDLINE=97000351; PubMed=8843436;
 RX Redenbach M., Kieser H.M., Denapait D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL109989; CAB53435.1;
 DR InterPro; IPR000410;
 DR InterPro; IPR003594;
 DR SMART; SM00387; HATPase_C; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 238 AA; 24931 MW; A927C944E16B34AF CRC64;

Query Match 14.3%; Score 86; DB 2; Length 238;
 Best Local Similarity 35.1%; Pred. No. 0.18;
 Matches 39; Conservative 8; Mismatches 48; Indels 16; Gaps 6;
 QY 9 RAAGARGCRLRSQVLPVVRALGLGRSDELVRFCSGSCRR--ARSPHDLSLASLLGAGA 66
 Db | | | | | : | | | | | : | | | | | : | | | | |
 2 RGARARGQAGGCHRLPVG---GRRPDHVRPRPGGAGSARDGIARDVHDLVIQRLF-AGA 56
 QY 67 LRPPEG-SRPVSQPCRCRPTRYEAVSFMD-----VNST---WRTVDRLSATA 108
 Db | | | | | : | | | | | : | | | | | : | | | | |
 57 LSPQALGRVTGRPKASERIQRVADLDDTIKVIIRSTIHALRESDRQTGTA 107

Search completed: August 16, 2001, 15:40:47
 Job time: 137 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:39:40 ; Search time 32.99 Seconds
(without alignments)
207.654 Million cell updates/sec

Title: us-09-357-349-3
Perfect score: 113
Sequence: 1 AGPGSRARAGARGCRLRS.....VNSTWRTVRLSATACGCLG 113

Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 113 | 100.0 | 113 | 21 | AA1984586 |
| 2 | 113 | 100.0 | 113 | 21 | AA198713 |
| 3 | 113 | 100.0 | 116 | 21 | AA1984587 |
| 4 | 113 | 100.0 | 116 | 21 | AA198712 |
| 5 | 113 | 100.0 | 139 | 21 | AA195011 |
| 6 | 113 | 100.0 | 140 | 21 | AA1984588 |
| 7 | 113 | 100.0 | 140 | 21 | AA198711 |
| 8 | 113 | 100.0 | 159 | 21 | AA194774 |
| 9 | 113 | 100.0 | 220 | 21 | AA1984583 |
| 10 | 113 | 100.0 | 220 | 21 | AA194776 |
| 11 | 113 | 100.0 | 220 | 21 | AA198710 |

| | | | | | |
|----|-----|-------|-----|----|-----------|
| 12 | 113 | 100.0 | 220 | 22 | AA1950978 |
| 13 | 113 | 100.0 | 228 | 21 | AA193559 |
| 14 | 113 | 100.0 | 228 | 21 | AA194775 |
| 15 | 113 | 100.0 | 229 | 21 | AA194582 |
| 16 | 113 | 100.0 | 237 | 21 | AA1984585 |
| 17 | 113 | 100.0 | 237 | 21 | AA192037 |
| 18 | 113 | 100.0 | 536 | 21 | AA1984592 |
| 19 | 100 | 88.5 | 113 | 21 | AA198709 |
| 20 | 100 | 88.5 | 116 | 21 | AA198708 |
| 21 | 100 | 88.5 | 140 | 21 | AA198707 |
| 22 | 100 | 88.5 | 237 | 21 | AA198706 |
| 23 | 96 | 85.0 | 96 | 21 | AA1984589 |
| 24 | 65 | 57.5 | 200 | 21 | AA198705 |
| 25 | 34 | 30.1 | 34 | 21 | AA198721 |
| 26 | 30 | 26.5 | 160 | 21 | AA1984595 |
| 27 | 29 | 25.7 | 29 | 21 | AA198722 |
| 28 | 28 | 24.8 | 28 | 21 | AA198720 |
| 29 | 26 | 23.0 | 123 | 21 | AA193665 |
| 30 | 26 | 23.0 | 222 | 21 | AA193560 |
| 31 | 26 | 23.0 | 224 | 21 | AA193558 |
| 32 | 26 | 23.0 | 224 | 21 | AA1984584 |
| 33 | 26 | 23.0 | 224 | 21 | AA198714 |
| 34 | 17 | 15.0 | 17 | 21 | AA198723 |
| 35 | 15 | 13.3 | 15 | 21 | AA198724 |
| 36 | 14 | 12.4 | 14 | 21 | AA198715 |
| 37 | 14 | 12.4 | 14 | 21 | AA198716 |
| 38 | 14 | 12.4 | 14 | 21 | AA198718 |
| 39 | 14 | 12.4 | 14 | 21 | AA198719 |
| 40 | 12 | 10.6 | 12 | 21 | AA198717 |
| 41 | 8 | 7.1 | 89 | 20 | AA196733 |
| 42 | 8 | 7.1 | 91 | 20 | AA196732 |
| 43 | 8 | 7.1 | 96 | 20 | AA196731 |
| 44 | 8 | 7.1 | 156 | 20 | AA196727 |
| 45 | 8 | 7.1 | 156 | 21 | AA192038 |

ALIGNMENTS

RESULT 1

AA1984586
ID AA1984586 standard; Protein; 113 AA.

XX AA1984586;

XX 25-JUL-2000 (first entry)

XX A first predicted human mature artemin polypeptide.

XX Human; artemin; growth factor; neurotrophic factor; trophic support;
KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
KW Parkinson's disease; Huntington's disease; acute brain injury;
KW acute spinal cord injury; nervous system tumour; blastoma;
KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
KW Parkinson's disease; small cell lung carcinoma.

OS Homo sapiens.

XX WO200018799-A1.

XX 06-APR-2000.

XX 29-SEP-1999; 99WO-US22604.

XX 29-SEP-1998; 98US-0163283.

XX 12-NOV-1998; 98US-0108148.

XX 22-DEC-1998; 98US-0218698.

XX (UNIW) UNIV WASHINGTON.

XX Milbrandt JD, Baloh RH;

KW Human; artemin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.
 XX Homo sapiens.
 OS W0200018799-A1.
 XX W0200018799-A1.
 XX 06-APR-2000.
 XX 29-SEP-1999; 99WO-US22604.
 XX 29-SEP-1998; 98US-0163283.
 PR 12-NOV-1998; 98US-0108148.
 PR 22-DEC-1998; 98US-0218698.
 XX (UNIW) UNIV WASHINGTON.
 PA Milbrandt JD, Balch RH;
 PI WPI: 2000-293109/25.
 DR N-PSDB; AAA12545.
 XX Isolated artemin growth factor proteins and the nucleic acids that
 PT encode them, useful for treating a range of degenerative neuronal
 PT disorders such as Parkinson's disease and Huntington's disease -
 XX Claim 4; Fig 3B; 96pp; English.

XX The present sequence represents a predicted mature human artemin growth
 CC factor protein. Artemin is a neurotrophic factor that belongs to the
 CC GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin
 CC family of growth factors and promotes differentiation, maintains mature
 CC phenotype and provides trophic support, promoting growth and survival of
 CC neurons. Artemin promotes the survival of trigeminal ganglion neurons,
 CC nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin
 CC is the only member of the GDNF family that binds to GFR-alpha (growth
 CC factor receptor-alpha) and activates the GFR-alpha/RET (ret protein-
 CC tyrosine kinase) receptor complex and additionally, like GDNF and
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin
 CC polypeptides and polynucleotides are administered to treat peripheral
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.
 CC idiopathic constipation or constipation associated with Parkinson's
 CC disease, spinal cord injury or use of opiate pain killers). They may
 CC also be used to treat a patient suffering from small cell lung carcinoma.
 XX Sequence 116 AA;

Query Match 100.0%; Score 113; DB 21; Length 116;
 Best Local Similarity 100.0%; Pred. No. 2.6e-99;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGGPGSRAAGARGCRLRSQVPRALGLHRSDELVFRFCSCRRARSPHDLAS 60
 Db 4 agpgsraaagargcrlrsqivpralglhrrsdelvfrfcsccrrarsphdlas 63
 QY 61 LLGAGALRPPGSRPVSQCRRPTREAVSFMDVNSTWRTVRLSATACGCG 113
 Db 64 llgagalrppgsrpsvqcrrptreavsfmdvnsrtvrlsatatcgclg 116

RESULT 4
 AAY68712

ID AAY68712 standard; Protein; 116 AA.
 AC AAY68712;
 XX 05-MAY-2000 (first entry)
 DE Amino acid sequence of a neublastin neurotrophic factor variant NBN116.
 KW Neurotrophic factor; neublastin; neurodegenerative disease;
 KW cerebral ischemic neuronal damage; traumatic brain injury;
 KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;
 KW Parkinson's disease; Parkinson-plus syndrome;
 KW progressive supranuclear palsy; Olivopontocerebellar atrophy;
 KW Shy-Drager Syndrome; Guamanian parkinsonism dementia complex;
 KW amyotrophic lateral sclerosis; memory impairment; neuronal disorder;
 KW neuropathy; ischemic stroke; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; multiple sclerosis;
 KW neurotoxin exposure; metabolic disease; diabetes; renal dysfunction;
 KW eye disorder.
 XX Homo sapiens.
 OS XX
 FH Key Location/Qualifiers
 FT Modified-site 98 /note= "glycosylated residue"
 FT XX
 PN W0200001815-A2.
 PD 13-JAN-2000.
 XX 05-JUL-1999; 99WO-DK00384.
 PR 06-JUL-1998; 98DK-0000904.
 PR 09-JUL-1998; 98US-0092229.
 PR 19-AUG-1998; 98DK-0001048.
 PR 25-AUG-1998; 98US-0097774.
 PR 06-OCT-1998; 98DK-0001265.
 PR 13-OCT-1998; 98US-0103908.
 PR 02-JUL-1999; 99US-0347613.
 PA (NEUR-) NEUROSEARCH AS.
 PI Johansen TE, Blom N, Hansen C;
 XX WPI: 2000-171013/15.
 DR New isolated polypeptides, used for treating e.g. neurodegenerative
 PT disease or disorder, neuronal damage or neuronal disorder of the
 PT peripheral nervous system, the medulla or the spinal cord -
 XX Claim 14; Page 98-99; 106pp; English.
 PS The present sequence represents a variant of a neurotrophic factor
 XX designated neublastin. Neublastin is a member of the glial cell line-
 XX derived neurotrophic factor sub-class of the transforming growth factor-
 XX beta superfamily of neurotrophic factors. Neublastin exhibits high
 XX affinity for the GFR-alpha3-RET receptor complex. The polypeptides can
 XX be used for treating a neurodegenerative disease or disorder, cerebral
 XX ischemic neuronal damage, traumatic brain injury, peripheral neuropathy,
 XX Alzheimer's disease, Huntington's disease, Parkinson's disease,
 XX Parkinson-plus syndromes, progressive supranuclear palsy,
 XX Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian
 XX parkinsonism dementia complex, amyotrophic lateral sclerosis, memory
 XX impairment, or a neuronal disorder of the peripheral nervous system,
 XX the medulla or the spinal cord. They can also be used for treating
 XX various neuropathies. They can also be used for treating ischemic stroke,
 XX acute brain injury, acute spinal cord injury, nervous system tumours,
 XX multiple sclerosis, exposure to neurotoxins, metabolic diseases such as
 XX diabetes or renal dysfunctions and damage caused by infectious agents,
 XX or various disorders in the eye.
 SQ Sequence 116 AA;

Query Match 100.0%; Score 113; DB 21; Length 116;
 Best Local Similarity 100.0%; Pred. No. 2.6e-99;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRARAAGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLISLAS 60
 |||||
 DB 4 agpgsraraagcrlrsqvlpralglghrsdvlrfrfcsgrarsphdlislas 63
 |||||
 QY 61 LLGAGALRPPGSRVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 |||||
 DB 64 llgagalrppgsvrpsqccrptryeavsfmdvnsdwtvdrisatagclg 116

RESULT 5

AAV45011
 ID AAV45011 standard; Protein; 139 AA.

AC AAV45011;

DT 17-MAY-2000 (first entry)

DE Partial human enovin protein.

KW Enovin; EVN; neurotrophic growth factor; chromosome lp31.3-32;
 KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;
 KW GDNF family receptor alpha-3; GFR alpha 3; neurotropic; analgesic;
 KW antirheumatic; cerebroprotective; antiparkinsonian; antinflammatory;
 KW antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;
 KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;
 KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;
 KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;
 KW Crohn's disease; bowel hypersensitivity.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..26

FT /label= Pro-sequence

FT /note= "Partial pro sequence of enovin"

FT Protein 27..139

FT /label= Mature_Enovin

FT /note= "Homologous to GDNF, Neurturin and Persephin"

FT Modified-site 121..123

FT /note= "Asn is N-glycosylated"

XX WO200004050-A2.

XX 27-JAN-2000.

XX 14-JUL-1999; 99WO-EP05031.

XX 14-JUL-1998; 98GB-0015283.

XX 12-FEB-1999; 99US-0248772.

XX 08-JUN-1999; 99US-0327668.

XX (JANC) JANSSEN PHARM NV.

XX Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;

XX WPI; 2000-182404/16.

XX N-PSDB; AA250705.

XX Novel human neurotrophic growth factor designated enovin used to treat
 PT neurological disorders, neuronal disorders, peripheral neuropathy,
 PT brain injury, nervous system disorders, prion associated and
 PT gastrointestinal diseases -

XX Claim 1; Fig 1; 125pp; English.

XX The present protein sequence is that of human enovin comprising

XX partial pro sequence. Enovin (EVN) is a neurotrophic growth

XX factor, that belongs to glial cell-line derived neurotrophic

CC factor (GDNF) family. It binds to GDNF family receptor
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome lp31.3-32.
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and
 CC prostate. It has neurotropic, analgesic, neuroprotective, antirheumatic,
 CC cerebroprotective, antiparkinsonian, antinflammatory, antidiarrhoeal,
 CC laxative and antiemetic activity. It can be used to treat neurological
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
 XX
 XX Sequence 139 AA;

Query Match 100.0%; Score 113; DB 21; Length 139;

Best Local Similarity 100.0%; Pred. No. 3e-99;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRARAAGCRLRSOLVPRALGLGHRSDLVRFRCGSCRRARSPHDLISLAS 60
 |||||
 DB 27 agpgsraraagcrlrsqvlpralglghrsdvlrfrfcsgrarsphdlislas 86
 |||||

QY 61 LLGAGALRPPGSRVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 |||||

DB 87 llgagalrppgsvrpsqccrptryeavsfmdvnsdwtvdrisatagclg 139

RESULT 6

AAV84588

ID AAV84588 standard; Protein; 140 AA.

XX AC AAV84588;

DT 25-JUL-2000 (first entry)

XX A third predicted human mature artemin polypeptide.

XX Human; artemin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.

XX Homo sapiens.

XX WO200018799-A1.

XX 06-APR-2000.

XX 29-SEP-1999; 99WO-US22604.

XX 29-SEP-1998; 98US-0163283.

XX 12-NOV-1998; 98US-0108148.

XX 22-DEC-1998; 98US-0218698.

XX (UNIW) UNIV WASHINGTON.

XX Milbrandt JD, Baloh RH;

XX WPI; 2000-293109/25.

XX N-PSDB; AAA12546.

XX Isolated artemin growth factor proteins and the nucleic acids that
 PT encode them, useful for treating a range of degenerative neuronal
 PT disorders such as Parkinson's disease and Huntington's disease -

XX Claim 4; Fig 3C; 96pp; English.

XX The present sequence represents a predicted mature human artemin growth

factor protein. Artemin is a neurotrophic factor that belongs to the GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin family of growth factors and promotes differentiation, maintains mature phenotype and provides trophic support, promoting growth and survival of neurons. Artemin promotes the survival of trigeminal ganglion neurons, nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin is the only member of the GDNF family that binds to GFR-alpha (growth factor receptor-alpha) and activates the GFR-alpha3/RET (ret protein-tyrosine kinase) receptor complex and additionally, like GDNF and neurturin, artemin also binds to and activates GFRalpha/RET. Artemin polypeptides and polynucleotides are administered to treat peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, acute brain injury, acute spinal cord injury, a nervous system tumour (e.g. blastomas), multiple sclerosis, infection or enteric disease (e.g. idiopathic constipation or constipation associated with Parkinson's disease, spinal cord injury or use of opiate pain killers). They may also be used to treat a patient suffering from small cell lung carcinoma.

Sequence 140 AA;

Query Match 100.0%; Score 113; DB 21; Length 140;
Best Local Similarity 100.0%; Pred. No. 3.1e-99;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSQVLPVRLGLGHRSDDELVRFCGSCRRARSPHDLAS 60
|||||
Db 28 agpgsraagaargcrlrsqvlvpralglghrsdelvrfcsgcrrarsphdlas 87
|||||

QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
|||||
Db 88 llgagalrpppgsrpvspccrptryeavsfmdvnstwtvdrisatagcglg 140
|||||

RESULT 7

AA168711
ID AAY68711 standard; Protein; 140 AA.

AC AAY68711;

DT 05-MAY-2000 (first entry)

DE Amino acid sequence of a neublastin neurotrophic factor variant NBN140.

KW Neurotrophic factor; neublastin; neurodegenerative disease;
KW cerebral ischemic neuronal damage; traumatic brain injury;
KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;
KW Parkinson's disease; Parkinson-plus syndrome;
KW progressive Supranuclear Palsy; Olivopontocerebellar atrophy;
KW amyotrophic lateral sclerosis; memory impairment; neuronal disorder;
KW neurotrophic lateral sclerosis; memory impairment; neuronal disorder;
KW acute spinal cord injury; acute brain injury;
KW neurotoxin exposure; metabolic disease; diabetes; multiple sclerosis;
KW eye disorder.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Modified-site 122 /note= "glycosylated residue"

XX W0200001815-A2.

PN 13-JAN-2000.

XX 05-JUL-1999; 99WO-DK00384.

XX 06-JUL-1998; 98DK-0000904.

PR 09-JUL-1998; 98US-0092229.

PR 19-AUG-1998; 98DK-0001048.

PR 25-AUG-1998; 98US-0097774.
PR 06-OCT-1998; 98DK-0001265.
PR 13-OCT-1998; 98US-0103908.
XX 02-JUL-1999; 99US-0347613.
XX (NEUR-) NEUROSEARCH AS.

XX Johansen TE, Blom N, Hansen C;

XX WPI; 2000-171013/15.

XX New isolated polypeptides, used for treating e.g. neurodegenerative disease or disorder, neuronal damage or neuronal disorder of the peripheral nervous system, the medulla or the spinal cord -
Claim 14; Page 98; 106pp; English.

CC The present sequence represents a variant of a neurotrophic factor designated neublastin. Neublastin is a member of the glial cell line-derived neurotrophic factor sub-class of the transforming growth factor-beta superfamily of neurotrophic factors. Neublastin exhibits high affinity for the GFR-alpha3-RET receptor complex. The polypeptides can be used for treating a neurodegenerative disease or disorder, cerebral ischemic neuronal damage, traumatic brain injury, peripheral neuropathy, Alzheimer's disease, Huntington's disease, Parkinson's disease, Parkinson-plus syndromes, progressive Supranuclear Palsy, Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian parkinsonism dementia complex, amyotrophic lateral sclerosis, memory impairment, or a neuronal disorder of the peripheral nervous system, the medulla or the spinal cord. They can also be used for treating various neuropathies. They can also be used for treating ischemic stroke, acute brain injury, acute spinal cord injury, nervous system tumours, multiple sclerosis, exposure to neurotoxins, metabolic diseases such as diabetes or renal dysfunctions and damage caused by infectious agents, or various disorders in the eye.

SQ Sequence 140 AA;

Query Match 100.0%; Score 113; DB 21; Length 140;
Best Local Similarity 100.0%; Pred. No. 3.1e-99;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLRSQVLPVRLGLGHRSDDELVRFCGSCRRARSPHDLAS 60
|||||
Db 28 agpgsraagaargcrlrsqvlvpralglghrsdelvrfcsgcrrarsphdlas 87
|||||

QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
|||||
Db 88 llgagalrpppgsrpvspccrptryeavsfmdvnstwtvdrisatagcglg 140
|||||

RESULT 8

AA144774

ID AAY44774 standard; Protein; 159 AA.

XX AAY44774;

XX 17-MAY-2000 (first entry)

XX Protein-2 comprising mature human Enovin sequence.

DE Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;

KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;
KW GDNF family receptor alpha-3; GFR alpha 3; neurotrophic; analgesic;
KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;
KW antidarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;
KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;
KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;
KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;
KW Crohn's disease; bowel hypersensitivity.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT Peptide 6..46
 FT /label= Pro-sequence
 FT /note= "Partial pro sequence of enovin"
 FT Protein 47..159
 FT /label= Mature Enovin
 FT /note= "Homologous to GDNF, Neurturin and Persephin"
 FT Modified-site 141..143
 FT /note= "Asn is N-glycosylated"
 FT Misc-difference 62
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 89
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 93
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 126
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 127
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 155
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 157
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT WO200004050-A2.
 PN
 XX
 XX 27-JAN-2000.
 XX
 XX 14-JUL-1999; 99WO-EP05031.
 PR 14-JUL-1998; 98GB-0015283.
 PR 12-FEB-1999; 99US-0248772.
 PR 08-JUN-1999; 99US-0327668.
 XX
 XX (JANC) JANSSEN PHARM NV.
 PA
 XX
 PI Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;
 DR WPI: 2000-182404/16.
 DR N-PSDB; AA250090.
 XX
 PT Novel human neurotrophic growth factor designated enovin used to treat neurological disorders, neuronal disorders, peripheral neuropathy, brain injury, nervous system disorders, prion associated and gastrointestinal diseases -
 PT
 XX
 PS Disclosure: Fig 3; 125pp; English.
 XX
 CC The present sequence is protein-2, comprising of partial pro sequence and mature human enovin sequence. Enovin (EVN) is a neurotrophic growth factor, that belongs to glial cell-line derived neurotrophic factor (GDNF) family. It binds to GDNF family receptor alpha-3 (GFR alpha 3). Enovin gene is located on chromosome lp31.3-32. It is predominantly expressed in heart, skeletal muscle, pancreas and prostate. It has nootropic, analgesic, neuroprotective, antirheumatic, cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal, laxative and antiemetic activity. It can be used to treat neurological disorders like Parkinson's, Alzheimer's and Huntington's disease, neuropathy, multiple sclerosis, endocrine neoplasia, prion associated diseases, stroke, pain, rheumatic/inflammatory diseases and gastrointestinal disorders like dyspepsia, constipation, intestinal atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
 XX Sequence 159 AA:

Query Match 100.0%; Score 113; DB 21; Length 159;
 Best Local Similarity 100.0%; Pred. No. 3,4e-99;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRARAAGARGCLRSQVVPVRLGLGHRSDLVRFRCGSCRRARSPHDLISLAS 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 47 agpgsraraagargclrsqvlvpralgighrsdelvrfrcsgscrrarsphdlislas 106
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATAGCGLG 113
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 107 llgagallrpppgsrpvspccrptryeavsfmdvnsstwtvdrslsatacgcig 159
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 9

AA184583
 ID AAY84583 standard; Protein; 220 AA.

XX
 AC AAY84583;

XX
 DT 25-JUL-2000 (first entry)

DE Amino acid sequence of a human pre-pro-artemin polypeptide.

XX Human; artemin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.

XX OS Homo sapiens.

XX WO200018799-A1.

XX PD 06-APR-2000.

XX 29-SEP-1999; 99WO-US22604.

XX 29-SEP-1998; 98US-0163283.

XX 12-NOV-1998; 98US-0108148.

XX 22-DEC-1998; 98US-0218698.

XX (UNIW) UNIV WASHINGTON.

XX Milbrandt JD, Baloh RH;

XX WPI: 2000-293109/25.

XX N-PSDB; AAA12540.

XX Isolated artemin growth factor proteins and the nucleic acids that encode them, useful for treating a range of degenerative neuronal disorders such as Parkinson's disease and Huntington's disease -

XX Claim 5; Fig 1B; 96pp; English.

XX The present sequence represents a pre-pro- artemin growth factor protein.
 CC Artemin is a neurotrophic factor that belongs to the GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin family of growth factors and promotes differentiation, maintains mature phenotype and provides trophic support, promoting growth and survival of neurons.
 CC Artemin promotes the survival of trigeminal ganglion neurons, nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin is the only member of the GDNF family that binds to GFR-alpha (growth factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-tyrosine kinase) receptor complex and additionally, like GDNF and neurturin, artemin also binds to and activates GFRalpha/RET. Artemin polypeptides and polynucleotides are administered to treat peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, acute brain

CC hypothalamic and other glandular, macrophagal, epithelial, stromal and
 CC blastocoeic disorders, and inflammatory, angiogenic and immunological
 CC disorders.
 XX
 SQ Sequence 220 AA;

Query Match 100.0%; Score 113; DB 22; Length 220;
 Best Local Similarity 100.0%; Pred. No. 4.5e-99;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLSOLVPRALGLGHRSDDELVRFCGSCRRARSPHDLISLAS 60
 |||||
 Db 108 agpgsraagargcrlsolvpralglghrsdelvrfcgscrrarsphdlislas 167

QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 |||||
 Db 168 llgagalrpppgsrpvspqccprtryeavsfmdvnstwtvdrisatagcglg 220

RESULT 13
 AAY93559
 ID AAY93559 standard; Protein: 228 AA.

XX AAY93559;

DT 25-SEP-2000 (first entry)

DE A human GDNF-related neurotrophic factor 4 (GRNF4).

KW GDNF; glial cell line-derived neurotrophic factor; GFRalpha-3;
 KW GDNF-related neurotrophic factor 4; GRNF4; GDNF family receptor-alpha-3;
 KW Parkinson's disease; Alzheimer's disease; anyotrophic lateral sclerosis;
 KW incontinence; bone loss; osteoporosis; osteogenesis imperfecta;
 KW hypercalcemia; nerve damage; stroke; cancer; dideoxycytidine; AIDS;
 KW chronic metabolic disease; renal dysfunction.

OS Homo sapiens.

XX WO200034475-A2.

PN 15-JUN-2000.

PD 08-DEC-1999; 99WO-US28975.

PE 09-DEC-1998; 98US-O111626.

PR (AMGE-) AMGEN INC.

PI Simonet WS, Asuncion FJ, Min H, Jing S;

XX WPI; 2000-423421/36.

DR N-PSDB; AAA46615.

XX New glial cell line-derived neurotrophic factor-related neurotrophic
 PT factor 4 useful for treating neurodegenerative disease such as
 PT Parkinson's disease and for treating nerve damage caused by physical
 PT injury and other metabolic diseases

XX Claim 1; Fig 7; 135pp; English.

XX The present sequence represents a human GDNF (glial cell line-derived
 CC neurotrophic factor)-related neurotrophic factor 4 (GRNF4) protein.
 CC The GRNF4 polypeptide is capable of binding a GDNF family
 CC receptor-alpha-3 (GFRalpha-3). The GRNF4 polynucleotides may be
 CC used for in vitro GRNF4 protein production as well as in cell therapy
 CC or gene therapy applications. GRNF4 protein product may be used in
 CC treating, Parkinson's disease, Alzheimer's disease, anyotrophic
 CC lateral sclerosis, incontinence, diseases associated with bone loss
 CC (e.g. osteoporosis, osteogenesis imperfecta or hypercalcemia of
 CC malignancy). GRNF4 protein products may also be used in the treatment
 CC of nerve damage which may occur to one or more types of nerve cells by
 CC physical injury, which causes the degeneration of the axonal processes

CC and/or nerve cell bodies near the site of injury, temporary or permanent
 CC cessation of blood flow to parts of the nervous system, as in stroke,
 CC intentional or accidental exposure to neurotoxins, for e.g.
 CC chemotherapeutic agents for the treatment of cancer or dideoxycytidine
 CC for the treatment of AIDS, chronic metabolic diseases, including
 CC diabetes or renal dysfunction.

XX Sequence 228 AA;

Query Match 100.0%; Score 113; DB 21; Length 228;
 Best Local Similarity 100.0%; Pred. No. 4.7e-99;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAGARGCRLSOLVPRALGLGHRSDDELVRFCGSCRRARSPHDLISLAS 60
 |||||
 Db 116 agpgsraagargcrlsolvpralglghrsdelvrfcgscrrarsphdlislas 175

QY 61 LLGAGALRPPPGSRPVSPQCCPRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 |||||
 Db 176 llgagalrpppgsrpvspqccprtryeavsfmdvnstwtvdrisatagcglg 228

RESULT 14
 AAY44775
 ID AAY44775 standard; Protein: 228 AA.

XX AAY44775;

DT 17-MAY-2000 (first entry)

DE Long splice variant of human Enovin.

KW Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;
 KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;
 KW GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;
 KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;
 KW antidiarthoeal; laxative; antiemetic; neurological disorder; Parkinson's;
 KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;
 KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;
 KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;
 KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..47

FT Peptide /label= Signal_Peptide

FT Peptide 48..115

FT /label= Pro_Sequence

FT Misc-difference 89..228

FT /note= "This region has been claimed specifically"

FT Protein 116..228

FT /label= Mature_Enovin

FT Modified-site 210..212

FT /note= "Homologous to GDNF, Neurturin and Persephin"

FT /note= "Asn is N-glycosylated"

FT Misc-difference 131

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 158

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 162

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 195

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 196

FT /note= "Conserved residue characteristic of Transforming

FT Growth Factor-beta (TGF-beta) family"

FT Misc-difference 224

FT /note= "Conserved residue characteristic of Transforming
FT Growth Factor-beta (TGF-beta) family"
FT Misc-difference 226
FT /note= "Conserved residue characteristic of Transforming
FT Growth Factor-beta (TGF-beta) family"
XX

PN WO200004050-A2.

XX

XX 27-JAN-2000.

XX PF 14-JUL-1999; 99WO-EP05031.

XX PR 14-JUL-1998; 98GB-0015283.

XX PR 12-FEB-1999; 99US-0248772.

XX PR 08-JUN-1999; 99US-0327668.

XX PA (JANC) JANSSEN PHARM NV.

XX PI Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;

XX DR WPI: 2000-182404/16.

XX DR N-PSDB; AAZ50091.

XX Novel human neurotrophic growth factor designated enovin used to treat

XX neurological disorders, neuronal disorders, peripheral neuropathy,

XX brain injury, nervous system disorders, prion associated and

XX gastrointestinal diseases

XX Claim 11; Fig 23; 125pp; English.

XX The present sequence is a long splice variant of human Enovin (EVN). EVN
CC is a neurotrophic growth factor, that belongs to glial cell-line derived
CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor
CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.
CC It is predominantly expressed in heart, skeletal muscle, pancreas and
CC prostate. It has nootropic, analgesic, neuroprotective, antidiarrhoeal,
CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,
CC laxative and antiemetic activity. It can be used to treat neurological
CC disorders like Parkinson's, Alzheimer's and Huntington's disease,
CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated
CC diseases, stroke, pain, rheumatic/inflammatory diseases and
CC gastrointestinal disorders like dyspepsia, constipation, intestinal
CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
CC EVN polynucleotide can be used in gene therapy.

XX Sequence 228 AA;

Query Match 100.0%; Score 113; DB 21; Length 228;

Best Local Similarity 100.0%; Pred. No. 4.7e-99;

Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLISLAS 60

DB 116 adgpgsraaagcrlrsqvlpralglghrsdelvrfrcsgscrrarsphdlsias 175

QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113

DB 176 llgagalrpppgsrpvspqccrptryeavsfmdvnstwrtvdrslsatcgclg 228

RESULT 15

AY84582

XX ID AAY84582 standard; Protein; 229 AA.

XX AC AAY84582;

XX DT 25-JUL-2000 (first entry)

XX Amino acid sequence of a human artemin polypeptide.

XX Human; artemin; growth factor; neurotrophic factor; trophic support;

XX neuron; trigeminal ganglion neuron; nodose ganglion neuron;

KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
KW Parkinson's disease; Huntington's disease; acute brain injury;
KW acute spinal cord injury; nervous system tumour; blastoma;
KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
KW Parkinson's disease; small cell lung carcinoma.

OS Homo sapiens.

XX WO2000018799-A1.

XX PD 06-APR-2000.

XX PF 29-SEP-1999; 99WO-US22604.

XX PR 29-SEP-1998; 98US-0163283.

XX PR 12-NOV-1998; 98US-0108148.

XX PR 22-DEC-1998; 98US-0218698.

XX PA (UNIW) UNIV WASHINGTON.

XX PI Milbrandt JD, Baloh RH;

XX DR WPI: 2000-293109/25.

XX DR N-PSDB; AA112539.

XX Isolated artemin growth factor proteins and the nucleic acids that
XX encode them, useful for treating a range of degenerative neuronal
XX disorders such as Parkinson's disease and Huntington's disease -

XX Disclosure; Fig 1A; 96pp; English.

XX The present sequence represents a human artemin growth factor protein.
CC Artemin is a neurotrophic factor that belongs to the GDNF (glial cell
CC line-derived neurotrophic factor)/neurturin/persephin family of growth
CC factors and promotes differentiation, maintains mature phenotype and
CC provides trophic support, promoting growth and survival of neurons.
CC Artemin promotes the survival of trigeminal ganglion neurons and nodose
CC ganglion neurons, superior cervical ganglion neurons and tyrosine-
CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin
CC is the only member of the GDNF family that binds to GFR-alpha (growth
CC factor receptor-alpha) and activates the GFR-alpha3/RET (RET protein-
CC tyrosine kinase) receptor complex and additionally, like GDNF and
CC neurturin, artemin also binds to and activates GFRalpha1/RET. Artemin
CC polypeptides and polynucleotides are administered to treat peripheral
CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,
CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain
CC injury, acute spinal cord injury, a nervous system tumour (e.g.
CC blastoma), multiple sclerosis, infection or enteric disease (e.g.
CC idiopathic constipation or constipation associated with Parkinson's
CC disease, spinal cord injury or use of opiate pain killers). They may
CC also be used to treat a patient suffering from small cell lung carcinoma.

XX Sequence 229 AA;

Query Match 100.0%; Score 113; DB 21; Length 229;

Best Local Similarity 100.0%; Pred. No. 4.7e-99;

Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAAGCRLRSQVPRALGLGHRSDLVRFRCGSCRRARSPHDLISLAS 60

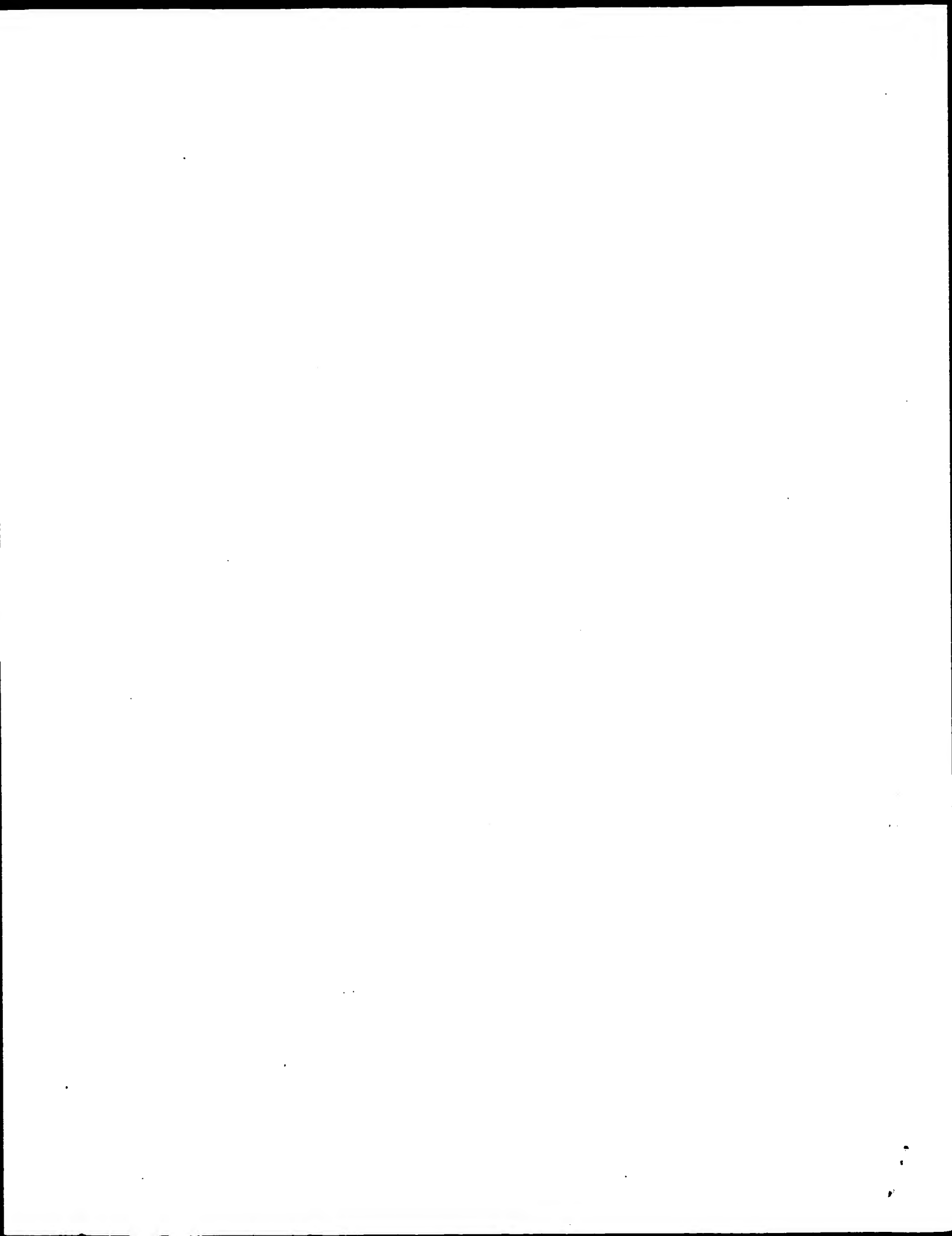
DB 103 agpgsraaagcrlrsqvlpralglghrsdelvrfrcsgscrrarsphdlsias 162

QY 61 LLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113

DB 163 llgagalrpppgsrpvspqccrptryeavsfmdvnstwrtvdrslsatcgclg 215

Search completed: August 16, 2001, 15:44:37

Job time: 297 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:40:50 ; Search time 18.89 Seconds
(without alignments)
123.171 Million cell updates/sec

Title: US-09-357-349-3
Perfect score: 113
Sequence: 1 AGGPGSRARAGRCRLRS.....VNSTWRTVDRLSATGCGLG 113

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|--------------------|
| 1 | 8 | 7.1 | 89 | US-08-931-858E-223 | Sequence 223, App |
| 2 | 8 | 7.1 | 96 | US-08-931-858E-221 | Sequence 221, App |
| 3 | 8 | 7.1 | 133 | US-08-931-858E-132 | Sequence 132, App |
| 4 | 8 | 7.1 | 156 | US-08-931-858E-217 | Sequence 217, App |
| 5 | 8 | 7.1 | 196 | US-08-481-814A-9 | Sequence 9, Appli |
| 6 | 8 | 7.1 | 365 | US-08-481-814A-10 | Sequence 10, Appli |
| 7 | 8 | 7.1 | 437 | US-08-136-119-4 | Sequence 4, Appli |
| 8 | 8 | 7.1 | 437 | US-07-882-711-2 | Sequence 2, Appli |
| 9 | 8 | 7.1 | 437 | US-08-723-415B-13 | Sequence 13, Appli |
| 10 | 8 | 7.1 | 437 | US-08-481-814A-6 | Sequence 6, Appli |
| 11 | 8 | 7.1 | 437 | US-08-462-174-2 | Sequence 2, Appli |
| 12 | 8 | 7.1 | 437 | US-08-801-092-1 | Sequence 1, Appli |
| 13 | 8 | 7.1 | 437 | US-09-189-627A-13 | Sequence 13, Appli |
| 14 | 8 | 7.1 | 476 | US-08-139-937-14 | Sequence 14, Appli |
| 15 | 8 | 7.1 | 476 | PCT-US93-11310-14 | Sequence 14, Appli |
| 16 | 7 | 6.2 | 7 | US-09-173-941-15 | Sequence 15, Appli |
| 17 | 7 | 6.2 | 13 | US-09-258-754-108 | Sequence 108, App |
| 18 | 7 | 6.2 | 13 | US-09-042-107-108 | Sequence 108, App |
| 19 | 7 | 6.2 | 94 | US-08-519-777-31 | Sequence 31, Appli |
| 20 | 7 | 6.2 | 94 | US-08-742-035-31 | Sequence 31, Appli |
| 21 | 7 | 6.2 | 94 | US-08-777-019-31 | Sequence 31, Appli |
| 22 | 7 | 6.2 | 94 | US-08-777-143-31 | Sequence 31, Appli |
| 23 | 7 | 6.2 | 94 | US-08-775-414-31 | Sequence 31, Appli |
| 24 | 7 | 6.2 | 94 | US-08-931-858E-31 | Sequence 31, Appli |
| 25 | 7 | 6.2 | 94 | US-08-981-739-31 | Sequence 31, Appli |
| 26 | 7 | 6.2 | 95 | US-08-931-858E-173 | Sequence 173, App |
| 27 | 7 | 6.2 | 95 | US-08-981-739-173 | Sequence 173, App |

| | | | | | | |
|----|---|-----|-----|---|-------------------|-------------------|
| 28 | 7 | 6.2 | 97 | 1 | US-08-443-568B-18 | Sequence 18, Appl |
| 29 | 7 | 6.2 | 97 | 5 | PCT-US94-06997-18 | Sequence 18, Appl |
| 30 | 7 | 6.2 | 102 | 1 | US-08-519-777-1 | Sequence 1, Appli |
| 31 | 7 | 6.2 | 102 | 1 | US-08-742-035-1 | Sequence 1, Appli |
| 32 | 7 | 6.2 | 102 | 2 | US-08-777-019-1 | Sequence 1, Appli |
| 33 | 7 | 6.2 | 102 | 2 | US-08-777-143-1 | Sequence 1, Appli |
| 34 | 7 | 6.2 | 102 | 3 | US-09-106-486-1 | Sequence 1, Appli |
| 35 | 7 | 6.2 | 102 | 3 | US-08-775-414-1 | Sequence 1, Appli |
| 36 | 7 | 6.2 | 102 | 4 | US-08-931-858E-1 | Sequence 1, Appli |
| 37 | 7 | 6.2 | 102 | 4 | US-08-981-739-1 | Sequence 1, Appli |
| 38 | 7 | 6.2 | 103 | 3 | US-09-106-486-4 | Sequence 4, Appli |
| 39 | 7 | 6.2 | 104 | 3 | US-08-775-414-79 | Sequence 79, Appl |
| 40 | 7 | 6.2 | 144 | 3 | US-08-775-414-81 | Sequence 81, Appl |
| 41 | 7 | 6.2 | 152 | 3 | US-08-775-414-83 | Sequence 83, Appl |
| 42 | 7 | 6.2 | 196 | 1 | US-08-063-552-7 | Sequence 7, Appli |
| 43 | 7 | 6.2 | 196 | 5 | PCT-US93-05704-7 | Sequence 7, Appli |
| 44 | 7 | 6.2 | 197 | 1 | US-08-519-777-7 | Sequence 7, Appli |
| 45 | 7 | 6.2 | 197 | 3 | US-09-106-486-5 | Sequence 5, Appli |

ALIGNMENTS

RESULT 1
US-08-931/858E-223
; Sequence 223, Application US/08931858E
; Patent No. 6222022
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, EUGENE M
; APPLICANT: MILLBRANDY, JEFFREY D
; APPLICANT: KOTZBAUER, PAUL T
; APPLICANT: LAMPE, PATRICIA A
; APPLICANT: KLEIN, ROBERT
; APPLICANT: DESAUVAGE, FRED
; TITLE-OF-INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patencin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,858E
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 971486
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 223:
SEQUENCE CHARACTERISTICS:
LENGTH: 89 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-931-858E-223

Query Match 7.1% Score 8; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 PCRPRTRY 86
Db 57 PCRPRTRY 64

RESULT 2
US-08-931-858E-221
; Sequence 221, Application US/08931858E
; Patent No. 6222022
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, EUGENE M
; APPLICANT: MILBRANDT, JEFFREY D
; APPLICANT: KOTZBAUER, PAUL T
; APPLICANT: LAMPE, PATRICIA A
; APPLICANT: KLEIN, ROBERT
; APPLICANT: DESAUVAGE, FRED
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931.858E
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 971486
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 221:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-931-858E-221

Query Match 7.1% Score 8; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 PCRPRTRY 86
Db 62 PCRPRTRY 69

RESULT 3
US-08-931-858E-132
; Sequence 132, Application US/08931858E
; Patent No. 6222022
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, EUGENE M
; APPLICANT: MILBRANDT, JEFFREY D
; APPLICANT: KOTZBAUER, PAUL T
; APPLICANT: LAMPE, PATRICIA A
; APPLICANT: KLEIN, ROBERT
; APPLICANT: DESAUVAGE, FRED
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
; NUMBER OF SEQUENCES: 239

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931.858E
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 971486
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-931-858E-132

Query Match 7.1% Score 8; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 PCRPRTRY 86
Db 99 PCRPRTRY 106

RESULT 4
US-08-931-858E-217
; Sequence 217, Application US/08931858E
; Patent No. 6222022
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, EUGENE M
; APPLICANT: MILBRANDT, JEFFREY D
; APPLICANT: KOTZBAUER, PAUL T
; APPLICANT: LAMPE, PATRICIA A
; APPLICANT: KLEIN, ROBERT
; APPLICANT: DESAUVAGE, FRED
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931.858E
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.

REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 971486
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 156 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-931-858E-217

Query Match 7.1%; Score 8; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 PCRPRTRY 86
|||||
DB 122 PCRPRTRY 129

RESULT 5
US-08-481-814A-9
Sequence 9, Application US/08481814A
Patent No. 5869040
GENERAL INFORMATION:
APPLICANT: Hsu, Yen-Ming
TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Biogen, Inc.
STREET: 14 Cambridge Center
CITY: Cambridge
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,814A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan, Warren A
REFERENCE/DOCKET NUMBER: A001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-679-2000
TELEFAX: 617-679-2838
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 196 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-814A-9

Query Match 7.1%; Score 8; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68
|||||
DB 19 LLGAGALR 26

RESULT 6
US-08-481-814A-10
Sequence 10, Application US/08481814A
Patent No. 5869040
GENERAL INFORMATION:
APPLICANT: Hsu, Yen-Ming
TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Biogen, Inc.
STREET: 14 Cambridge Center
CITY: Cambridge
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,814A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan, Warren A
REFERENCE/DOCKET NUMBER: A001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-679-2000
TELEFAX: 617-679-2838
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-814A-10

Query Match 7.1%; Score 8; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68
|||||
DB 19 LLGAGALR 26

RESULT 7
US-08-136-119-4
Sequence 4, Application US/08136119
Patent No. 5473056
GENERAL INFORMATION:
APPLICANT: Heimbros, David C.
APPLICANT: Hoyle, Mona I.
APPLICANT: Oliff, Allen I.
TITLE OF INVENTION: E 2F-2, A NOVEL MAMMALIAN TRANSCRIPTION
TITLE OF INVENTION: FACTOR
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Muthard
STREET: P.O. Box 2000, 126 Lincoln Avenue
CITY: Rahway
STATE: N.J.
COUNTRY: USA
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,119
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Muthard, David A.
REGISTRATION NUMBER: 35,297
REFERENCE/DOCKET NUMBER: 19042
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-3903
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-136-119-4

Query Match 7.1%; Score 8; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68
Db 19 LLGAGALR 26

RESULT 8
US-07-882-711-2
Sequence 2, Application US/07882711
Patent No. 5759803
GENERAL INFORMATION:
APPLICANT: Kaelin Jr., William G.
APPLICANT: Flemington, Erik
APPLICANT: Sellers, William
APPLICANT: DeCaprio, James A.
APPLICANT: Livingston, David M.
TITLE OF INVENTION: Retinoblastoma Associated Protein-1 cDNA
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Choate, Hall & Stewart
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/882,711
FILING DATE: 19920513
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kennedy, Bill
REGISTRATION NUMBER: 33,407
REFERENCE/DOCKET NUMBER: DFCI#236
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-5020
TELEFAX: (617) 227-7566
TELEX: 289374
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: AMINO ACID

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-07-882-711-2

Query Match 7.1%; Score 8; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68
Db 19 LLGAGALR 26

RESULT 9
US-08-723-415B-13
Sequence 13, Application US/08723415B
Patent No. 5859199
GENERAL INFORMATION:
APPLICANT: Lathangue, Nicholas B.
APPLICANT: delaluna, Susana
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 No. 5859199th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,415B
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9610195.1
FILING DATE: 15-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 117-220
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-723-415B-13

Query Match 7.1%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGALR 68
Db 19 LLGAGALR 26

RESULT 10
US-08-481-814A-6
Sequence 6, Application US/08481814A
Patent No. 5869040

GENERAL INFORMATION:
APPLICANT: Hsu, Yen-Ming
TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Blegen, Inc.
STREET: 14 Cambridge Center
CITY: Cambridge
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,814A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan, Warren A
REFERENCE/DOCKET NUMBER: A001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-679-2000
TELEFAX: 617-679-2838
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-814A-6

Query Match 7.1%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 LLAGALR 68
Db 19 LLAGALR 26

RESULT 11
US-08-462-174-2
Sequence 2, Application US/08462174
Patent No. 5981723
GENERAL INFORMATION:
APPLICANT: Kaelin Jr., William G.
APPLICANT: Flemington, Erik
APPLICANT: Sellers, William
APPLICANT: Decaprio, James A.
APPLICANT: Livingston, David M.
TITLE OF INVENTION: Retinoblastoma Associated Protein-1 cDNA
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Choate, Hall & Stewart
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,174

FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/882,711
FILING DATE: 13-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kennedy, Bill
REGISTRATION NUMBER: 33,407
REFERENCE/DOCKET NUMBER: DFCI#236
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-5020
TELEFAX: (617) 227-7566
TELEX: 289374
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-462-174-2

Query Match 7.1%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 LLAGALR 68
Db 19 LLAGALR 26

RESULT 12
US-08-801-092-1
Sequence 1, Application US/08801092
Patent No. 6074850
GENERAL INFORMATION:
APPLICANT: Antelman, Douglas
APPLICANT: Gregory, Richard J.
APPLICANT: Wills, Kenneth N.
TITLE OF INVENTION: Tissue Specific Expression of
TITLE OF INVENTION: Retinoblastoma Protein
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,092
FILING DATE: 14-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,517
FILING DATE: 15-NOV-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Filts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 016930-001020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 703-576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids

; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-801-092-1

Query Match 7.1%; Score 8; DB 3; Length 437;
Best Local Similarity 100.0%; Pred. No. 8.7; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0

QY 61 LLGAGALR 68
| | | | |
Db 19 LLGAGALR 26

RESULT 13
US-09-189-627A-13
; Sequence 13, Application US/09189627A
; Patent No. 6159691
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/189,627A
; CURRENT FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 437
; TYPE: PRT
; ORGANISM: human
US-09-189-627A-13

Query Match 7.1%; Score 8; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 8.7; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0

QY 61 LLGAGALR 68
| | | | |
Db 19 LLGAGALR 26

RESULT 14
US-08-139-937-14
; Sequence 14, Application US/08139937
; Patent No. 5821070
; GENERAL INFORMATION:
; APPLICANT: LEE, WEN-HWA
; APPLICANT: SHAN, BEI
; TITLE OF INVENTION: CELLULAR GENES ENCODING
; TITLE OF INVENTION: RETINOBLASTOMA-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/139,937

; FILING DATE: 20-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/979,156
; FILING DATE: 20-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-CJ 9370
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-139-937-14

Query Match 7.1%; Score 8; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 9.4; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0

QY 61 LLGAGALR 68
| | | | |
Db 58 LLGAGALR 65

RESULT 15
PCT-US93-11310-14
; Sequence 14, Application PC/TUS9311310
; GENERAL INFORMATION:
; APPLICANT: BOARD OF REGENTS OF THE UNIVERSITY OF TEXAS SYSTEM
; TITLE OF INVENTION: CELLULAR GENES ENCODING
; TITLE OF INVENTION: RETINOBLASTOMA-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11310
; FILING DATE: 19-NOV-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: FP-CJ 9790
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-11310-14

Query Match 7.1%; Score 8; DB 5; Length 476;
Best Local Similarity 100.0%; Pred. No. 9.4; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0

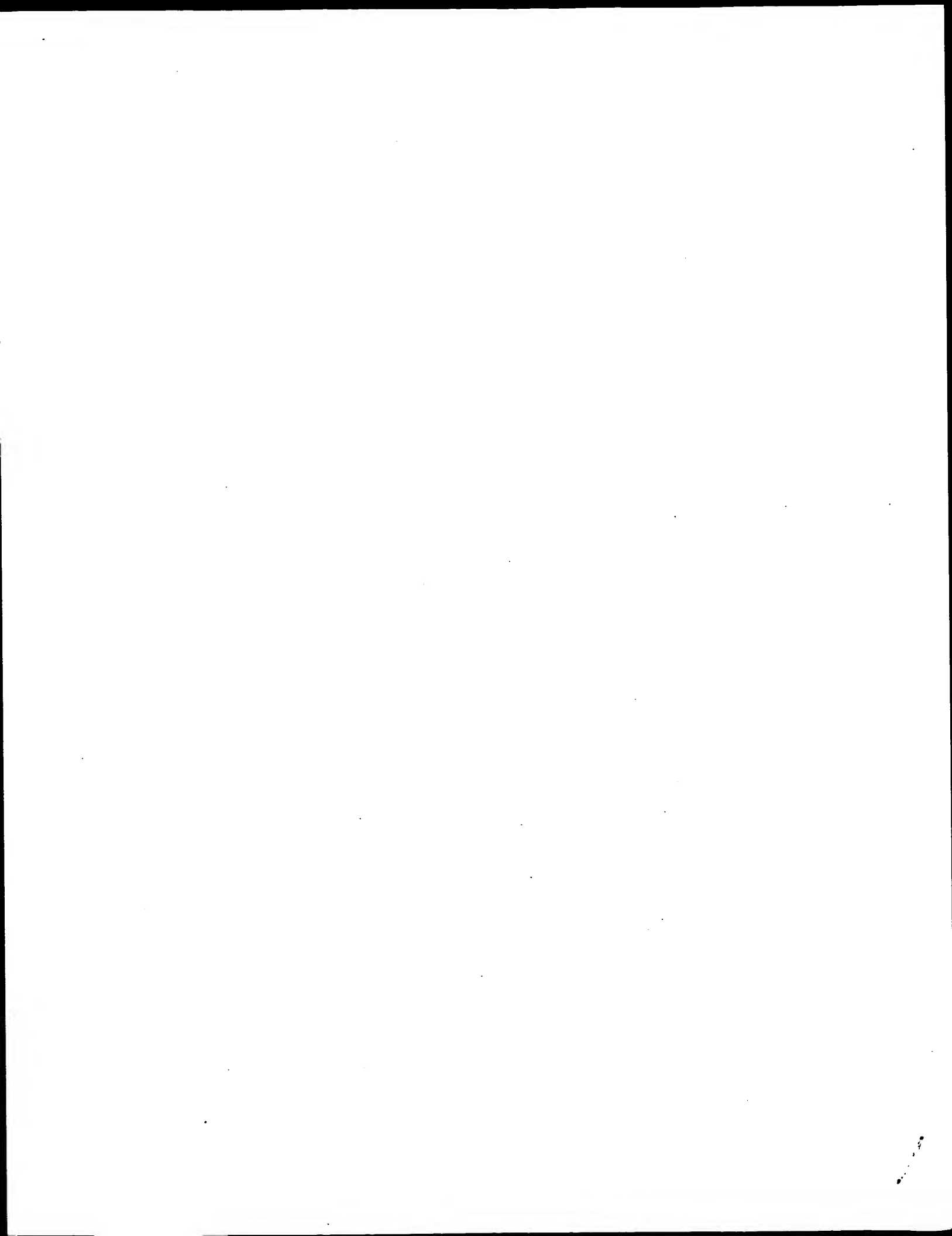
Fri Aug 17 09:56:20 2001

us-09-357-349-3.oli.ra

Page 7

Oy 61 LLGAGALR 68
| | | | |
Db 58 LLGAGALR 65

Search completed: August 16, 2001, 15:45:31
Job time: 281 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:40:10 ; Search time 21.79 Seconds
(without alignments)
395.031 Million cell updates/sec

Title: US-09-357-349-3
Perfect score: 113
Sequence: 1 AGPGSRAAGAGCRLRS.....VNSTWRTVDRLSATACGLG 113

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 8 | 7.1 | 108 | 2 S28241 | NADH dehydrogenase |
| 2 | 8 | 7.1 | 430 | 2 A56209 | transcription fact |
| 3 | 8 | 7.1 | 437 | 2 JC4929 | transcription fact |
| 4 | 7 | 6.2 | 133 | 2 E70507 | hypothetical prote |
| 5 | 7 | 6.2 | 133 | 2 I51960 | WTL zinc-finger ho |
| 6 | 7 | 6.2 | 138 | 2 PC1195 | genome polyprotein |
| 7 | 7 | 6.2 | 197 | 2 T47159 | hypothetical prote |
| 8 | 7 | 6.2 | 201 | 2 A47096 | response regulator |
| 9 | 7 | 6.2 | 210 | 2 E84231 | hypothetical prote |
| 10 | 7 | 6.2 | 213 | 2 D81057 | GDSL lipase family |
| 11 | 7 | 6.2 | 228 | 2 S67199 | hypothetical prote |
| 12 | 7 | 6.2 | 231 | 2 T76204 | hypothetical prote |
| 13 | 7 | 6.2 | 234 | 2 E72679 | hypothetical prote |
| 14 | 7 | 6.2 | 243 | 2 T34870 | probable membrane |
| 15 | 7 | 6.2 | 260 | 1 E64394 | hypothetical prote |
| 16 | 7 | 6.2 | 274 | 4 JC5585 | hypothetical mutan |
| 17 | 7 | 6.2 | 303 | 2 T35930 | hypothetical prote |
| 18 | 7 | 6.2 | 306 | 2 H75358 | LAG/AO transport s |
| 19 | 7 | 6.2 | 308 | 2 F72508 | probable hydrogena |
| 20 | 7 | 6.2 | 310 | 2 C83619 | probable acyl tran |
| 21 | 7 | 6.2 | 319 | 2 JC4390 | bone marrow stroma |
| 22 | 7 | 6.2 | 323 | 2 T35734 | probable aminopept |
| 23 | 7 | 6.2 | 339 | 2 B69436 | LSU ribosomal prot |
| 24 | 7 | 6.2 | 344 | 2 T34981 | probable integral |
| 25 | 7 | 6.2 | 381 | 2 S16506 | hypothetical prote |
| 26 | 7 | 6.2 | 394 | 2 T35447 | hypothetical prote |
| 27 | 7 | 6.2 | 396 | 1 YTEC32 | tetracycline resis |
| 28 | 7 | 6.2 | 402 | 2 A83398 | probable MFS trans |
| 29 | 7 | 6.2 | 404 | 2 F59097 | hypothetical prote |

30 7 6.2 410 2 JC5046 Wilms' tumor suppr
31 7 6.2 431 2 S09824 hypothetical prote
32 7 6.2 446 2 C64205 signal recognition
33 7 6.2 448 2 S33926 Wilms' tumor prote
34 7 6.2 449 2 S38080 Wilms' tumor suscep
35 7 6.2 449 2 A39692 Wilms' tumor prote
36 7 6.2 458 2 T19941 hypothetical prote
37 7 6.2 459 2 JH0594 vasoactive intesti
38 7 6.2 527 1 SAHUP saposin precursor
39 7 6.2 601 2 JC4576 serine proteinase
40 7 6.2 604 2 A69858 ABC transporter (A
41 7 6.2 857 2 F84227 molybdenum-binding
42 7 6.2 863 2 S37040 paramyosin - tapew
43 7 6.2 977 2 G72526 hypothetical prote
44 7 6.2 1008 2 T04462 hypothetical prote
45 7 6.2 1069 2 D85383 hypothetical prote

ALIGNMENTS

RESULT 1

S28241
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain CI-AGGG - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 22-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
C:Accession: S28241
R:Walker, J.E.; Arizmendi, J.M.; Dupuis, A.; Fearnley, I.M.; Finel, M.; Medd, S.M.; P
J. Mol. Biol. 226, 1051-1072, 1992
A:Title: Sequences of 20 subunits of NADH: ubiquinone oxidoreductase from bovine hear
A:Reference number: S28237; MUID:92389317
A:Accession: S28241
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-108 <WAL>
A:Cross-references: EMBL:X63216; NID:g231; PIDN:CAA44901.1; PID:g232
C:Keywords: NAD; oxidoreductase

Query Match 7.1%; Score 8; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 ARAAGARG 15
Db 26 ARAAGARG 33

RESULT 2

A56209
transcription factor E2F1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
C:Accession: A56209
R:Li, Y.; Slansky, J.E.; Myers, D.J.; Drinkwater, N.R.; Kaelin, W.G.; Farnham, P.J.
Mol. Cell. Biol. 14, 1861-1869, 1994
A:Title: Cloning, chromosomal location, and characterization of mouse E2F1.
A:Reference number: A56209; MUID:94158858
A:Accession: A56209
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-430 <LIA>
A:Cross-references: GB:L21973; NID:g1091479; PIDN:AAA83217.1; PID:g405524

Query Match 7.1%; Score 8; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 LLGAGALR 68
Db 17 LLGAGALR 24

```

RESULT 3
JC4929
N:transcription factor E2F1 - human
N:alternate names: Ap12 protein; retinoblastoma-associated protein; retinoblastoma-binding
C:Species: Homo sapiens (man)
C:Date: 22-Oct-1996 #sequence_revision 01-Nov-1996 #text_change 02-Jun-2000
C:Accession: JC4929; A45032; A42998; A42997; I54091
R:Neuman, E.; Sellers, W.R.; McNeil, J.A.; Lawrence, J.B.; Kaelin Jr., W.G.
Gene 173, 163-169, 1996
A:Title: Structure and partial genomic sequence of the human E2F1 gene.
A:Reference number: JC4929; MUID:97082961
A:Accession: JC4929
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-437 <NEU>
A:Cross-references: GB:U47675; NID:g1594281; GB:U47676; NID:g1594282; GB:U47677; NID:g1594283
A:Experimental source: placenta
R:Shan, B.; Zhu, X.; Chen, P.L.; Durfee, T.; Yang, Y.; Sharp, D.; Lee, W.H.
Mol. Cell. Biol. 12, 5620-5631, 1992
A:Title: Molecular cloning of cellular genes encoding retinoblastoma-associated proteins
A:Reference number: A45032; MUID:93078763
A:Accession: A45032
A:Molecule type: mRNA
A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <HEL>
A:Cross-references: GB:S49592; NID:g260573; PIDN:AAB24289.1; PID:g181918
A:Note: sequence extracted from NCB1 backbone (NCBIP:119095)
A:Note: the authors are uncertain whether Met-1 is the initiator or whether translation
R:Helin, K.; Lees, J.A.; Vidal, M.; Dyson, N.; Harlow, E.; Pattaeay, A.
Cell 70, 337-350, 1992
A:Title: A cDNA encoding a PRB-binding protein with properties of the transcription factor
A:Reference number: A42998; MUID:92346720
A:Accession: A42998
A:Molecule type: mRNA
A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <HEL>
A:Cross-references: GB:M6577; NID:g181917; PIDN:AAA35782.1; PID:g181918
A:Experimental source: Nalm 6 pre-B leukemia cell line
A:Note: sequence extracted from NCB1 backbone (NCBIN:110015, NCBIP:110016)
R:Kaelin Jr., W.G.; Krek, W.; Sellers, W.R.; DeCaprio, J.A.; Ajchenbaum, F.; Fuchs, C.S.
Cell 70, 351-364, 1992
A:Title: Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F
A:Reference number: A42997; MUID:92346721
A:Accession: A42997
A:Molecule type: mRNA
A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <KAE>
A:Cross-references: GB:U13851; NID:g595713
A:Experimental source: Akata cells; expression vector pGEX-2TK
A:Note: sequence extracted from NCB1 backbone (NCBIN:110018, NCBIP:110019)
R:Johnson, D.G.; Ontani, K.; Nevins, J.R.
Genes Dev. 8, 1514-1525, 1994
A:Title: Autoregulatory control of E2F1 expression in response to positive and negative
A:Reference number: I54091; MUID:95047311
A:Accession: I54091
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-88, 'R', 'T', 122-123, 'TPR', 127, 'QRR', 297-299, 'PRR', 308-309, 'RA', 312, 'C' <RES>
A:Cross-references: GB:S74230; NID:g712816; PIDN:AAD14150.1; PID:g4261850
C:Genetics:
A:Gene: GDB:E2F1
A:Cross-references: GDB:I34661; OMIM:189971
A:Map position: 20q11-20q11
A:Introns: 87/3 118/1 191/2 242/2 280/3 356/1
C:Keywords: DNA binding; transcription factor.
F:67-108/Region: cyclin box #status predicted
F:118-190/Domain: DNA binding #status predicted <DNA>
F:191-241/Region: 7-residue repeats

Query Match 7.1%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

~QY 61 LLGAGALR 68

Db 19 LLGAGALR 26
|||||||
RESULT 4
E70507
hypothetical protein RV2060 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: E70507
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: E70507
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-133 <COL>
A:Cross-references: GB:297984; GB:AL123456; NID:g3261833; PIDN:CAB10786.1; PID:e12999
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2060

Query Match 6.2%; Score 7; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 VVVRALG 29
|||||||
Db 35 VVVRALG 41

RESULT 5
I51960
WT1 zinc-finger homolog, testis - rat (fragment)
C:Species: Rattus sp. (rat)
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 05-Nov-1999
C:Accession: I51960
R:Madden, S.L.; Rauscher, F.J.
Ann. N. Y. Acad. Sci. 684, 75-84, 1993
A:Title: Positive and negative regulation of transcription and cell growth mediated b
A:Reference number: I51960; MUID:93304850
A:Accession: I51960
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-133 <RES>
A:Cross-references: GB:S63358; NID:g386513; PIDN:AAB27319.1; PID:g386514

Query Match 6.2%; Score 7; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 RSDELVR 39
|||||||
Db 111 RSDELVR 117

RESULT 6
PC1195
genome polyprotein - hepatitis C virus (strain NV4-2) (fragment)
C:Species: hepatitis C virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000
C:Accession: PC1195
R:Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Sekiya, H.; Hijikata, M.; Shimo
Biochem. Biophys. Res. Commun. 189, 119-127, 1992
A:Title: Characterization of hypervariable regions in the putative envelope protein o
A:Reference number: PC1182; MUID:93080545
A:Accession: PC1195

```

A;Status: nucleic acid sequence not shown

A:Molecule type: genomic RNA

A:Residues: 1-138 <KAT>

A:Cross-references: GB:D12955; NID:g285850; PIDN:BAR02331.1; PID:g285851

C:Superfamily: hepatitis C virus genome polypeptide

C:Keywords: polyprotein

Query Match 6.2%; Score 7; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 SLASLLG 63

Db 30 SLASLLG 36

RESULT 7

T47159

hypothetical protein DKFZp762B0211.1 - human

C:Species: Homo sapiens (man)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47159

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z4379

A:Accession: T47159

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-197 <NAA>

A:Cross-references: EMBL:AL161995

A:Experimental source: adult melanoma (MeWo cell line); clone DKFZp762B0211

C:Genetics:

A:Note: DKFZp762B0211.1

Query Match 6.2%; Score 7; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 OPCCRPT 84

Db 162 OPCCRPT 168

RESULT 8

A47096

response regulator homolog, aerial mycelium formation - Streptomyces griseus

C:Species: Streptomyces griseus

C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jul-2000

C:Accession: A47096

R:Ueda, K.; Miyake, K.; Horinouchi, S.; Beppu, T.

J. Bacteriol. 175, 2006-2016, 1993

A:Title: A gene cluster involved in aerial mycelium formation in Streptomyces griseus

A:Reference number: A47096; MUID:93209944

A:Accession: A47096

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-201 <UED>

A:Cross-references: GB:D13614; NID:g3721543; PIDN:BAA33536.1; PID:g303913

C:Superfamily: regulatory protein comA; response regulator homology

C:Keywords: DNA binding; phosphoprotein

F;54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 6.2%; Score 7; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 SLASLLG 63

Db 16 SLASLLG 22

RESULT 9

E84231

hypothetical protein Vng0740c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: E84231

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky

; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483

A:Accession: E84231

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-210 <STO>

A:Cross-references: GB:AE004437; NID:gl0580319; PIDN:AAG19217.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG0740C

Query Match 6.2%; Score 7; DB 2; Length 210;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 LLGAGAL 67

Db 167 LLGAGAL 173

RESULT 10

D81057

GD5L lipase family protein NMB1674 [imported] - Neisseria meningitidis (strain MC58 s

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001

C:Accession: D81057; B81821

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: D81057

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <TET>

A:Cross-references: GB:AE002517; GB:AE002098; NID:g7226912; PIDN:AAF42023.1; PID:g722

A:Experimental source: serogroup B, strain MC58

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491

A:Reference number: A81775; MUID:20222556

A:Accession: B81821

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85152.1; PID:g738

A:Experimental source: serogroup A, strain 22491

C:Genetics:

A:Gene: NMB1674; NMA1932

Db 13 LLGAGAL 19

RESULT 11

S67199

hypothetical protein YOR295w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein O5620

C:Species: Saccharomyces cerevisiae

C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000

C:Accession: S67199

R:Ziepluch, C.; Jauniaux, J.C.; Kordes, E.; Poirrey, R.; Pujol, A.; Tobiasch, E.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67194

A:Accession: S67199

A:Molecule type: DNA

A:Residues: 1-228 <CZ1>

A:Cross-references: EMBL:R75203; NID:gl420653; PIDN:CAA99523.1; PID:gl420654; GSPDB:GN00

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YOR295w

A:Map position: 15R

C:Superfamily: hypothetical protein YMR233w

Query Match

Best Local Similarity 6.2%; Score 7; DB 2; Length 228;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 SLASLLG 63

Db 129 SLASLLG 135

RESULT 12

S76204

hypothetical protein - Synecocystis sp. (strain PCC 6803)

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997

C:Accession: S76204

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

s.

A:Reference number: S74322; MUID:97061201

A:Accession: S76204

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-231 <KAN>

A:Cross-references: EMBL:D90914; GB:AB001339; NID:gl653477; PIDN:BAA18463.1; PID:gl65355

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Start codon: GTG

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

F:156-230/Domain: double-stranded RNA-binding repeat homology <DSR2>

Query Match

Best Local Similarity 6.2%; Score 7; DB 2; Length 231;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 RLRSQLV 23

Db 67 RLRSQLV 73

RESULT 13

E72679

hypothetical protein APE0857 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000

C:Accession: E72679

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J. DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: E72679

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-234 <KAN>

A:Cross-references: DDBJ:AF000060; NID:gs104188; PIDN:BAA79837.1; PID:dl043623; PID:g

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0857

C:Superfamily: Aeropyrum pernix hypothetical protein APE0857

Query Match

Best Local Similarity 6.2%; Score 7; DB 2; Length 234;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 RPPGSR 74

Db 207 RPPGSR 213

RESULT 14

T34870

probable membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 08-Sep-2000

C:Accession: T34870

R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21560

A:Accession: T34870

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-243 <SEE>

A:Cross-references: EMBL:AL109849; PIDN:CAB52862.1; GSPDB:GN00070; SCOEDB:SC3A3.05C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC3A3.05C.

C:Superfamily: Streptomyces coelicolor probable membrane protein SC3A3.05C

Query Match

Best Local Similarity 6.2%; Score 7; DB 2; Length 243;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLGAGAL 67

Db 181 LLGAGAL 187

RESULT 15

E64394

hypothetical protein MJ0757 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: E64394

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc

A:Reference number: A64300; MUID:96337999

A:Accession: E64394

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-260 <BUL>

A:Cross-references: GB:U67521; GB:L77117; NID:gl591463; PIDN:AAB98749.1; PID:gl591468

C:Genetics:

A:Map position: REV682205-681423
A:Start codon: TTG
C:Superfamily: hypothetical protein AFL307

Query Match 6.2%; Score 7; DB 1; Length 260;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 PHDLSLA 59
|||||||
DB 123 PHDLSLA 129

Search completed: August 16, 2001, 15:45:06
Job time: 296 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2001, 07:35:02 ; Search time 15.24 Seconds
(without alignments)
258.331 Million cell updates/sec

Title: US-09-357-349-3

Perfect score: 113

Sequence: 1 AGGPGSRARAGRCRLRS.....VNSTWRTVDRLSATACGCLG 113

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 94743 seqs, 34840360 residues
Word size : 0

Total number of hits satisfying chosen parameters: 94743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 8 | 7.1 | 108 | 1 NIGM_BOVIN | Q02374 bos taurus |
| 2 | 8 | 7.1 | 156 | 1 P2PN_HUMAN | O60542 homo sapien |
| 3 | 8 | 7.1 | 430 | 1 E2F1_MOUSE | Q61501 mus musculus |
| 4 | 8 | 7.1 | 437 | 1 E2F1_HUMAN | Q01094 homo sapien |
| 5 | 7 | 6.2 | 68 | 1 BD01_HUMAN | O09753 homo sapien |
| 6 | 7 | 6.2 | 68 | 1 BD01_MACMU | O18794 macaca mula |
| 7 | 7 | 6.2 | 197 | 1 NRTN_HUMAN | O99748 homo sapien |
| 8 | 7 | 6.2 | 207 | 1 UCRI_CHRVI | O31214 chromatium |
| 9 | 7 | 6.2 | 231 | 1 RNC_SYNV3 | P74368 synchocyst |
| 10 | 7 | 6.2 | 239 | 1 WT1_SMIMA | P49953 smnthopsis |
| 11 | 7 | 6.2 | 242 | 1 YBL1_STRCI | P33653 streptomyce |
| 12 | 7 | 6.2 | 260 | 1 Y757_METJA | Q58167 methanococc |
| 13 | 7 | 6.2 | 288 | 1 WT1_ALLMI | P50902 alligator m |
| 14 | 7 | 6.2 | 319 | 1 BST1_RAT | O63072 rattus norv |
| 15 | 7 | 6.2 | 323 | 1 PIP_STRCO | O95214 streptomyce |
| 16 | 7 | 6.2 | 339 | 1 RLA0_ARCFU | O28781 archaeoglob |
| 17 | 7 | 6.2 | 396 | 1 TCR3_ECOLI | P02981 escherichia |
| 18 | 7 | 6.2 | 431 | 1 UL61_HCMVA | P16818 human cytom |
| 19 | 7 | 6.2 | 446 | 1 SR54_MYCGE | P47294 mycoplasma |
| 20 | 7 | 6.2 | 448 | 1 WT1_RAT | P49952 rattus norv |
| 21 | 7 | 6.2 | 449 | 1 WT1_HUMAN | P19544 homo sapien |
| 22 | 7 | 6.2 | 449 | 1 WT1_MOUSE | P22561 mus musculus |
| 23 | 7 | 6.2 | 449 | 1 WT1_PIG | O62651 sus scrofa |
| 24 | 7 | 6.2 | 459 | 1 VIPR_RAT | P30083 rattus norv |
| 25 | 7 | 6.2 | 473 | 1 TED_HUMAN | O75949 homo sapien |
| 26 | 7 | 6.2 | 524 | 1 SAP_HUMAN | P07602 h proactiva |
| 27 | 7 | 6.2 | 863 | 1 MYSP_ECHGR | P35417 echinococcu |
| 28 | 7 | 6.2 | 863 | 1 MYSP_TAESO | P35418 taenia soli |
| 29 | 7 | 6.2 | 1069 | 1 S24B_ARATH | Q9m081 arabidopsis |
| 30 | 7 | 6.2 | 1075 | 1 NFC3_HUMAN | Q12968 homo sapien |
| 31 | 7 | 6.2 | 1075 | 1 NFC3_MOUSE | P97305 mus musculus |
| 32 | 7 | 6.2 | 1147 | 1 SREL_HUMAN | P36956 homo sapien |
| 33 | 6 | 5.3 | 79 | 1 Y437_METJA | Q37879 methanococc |

34 6 5.3 100 1 CHA3_BOMMO
35 6 5.3 115 1 NU3M_DASNO
36 6 5.3 116 1 NLTP_GERHY
37 6 5.3 116 1 ULC9_HCMVA
38 6 5.3 119 1 YN8E_YEAST
39 6 5.3 120 1 NLTP_DAUCA
40 6 5.3 120 1 R18E_AERPE
41 6 5.3 122 1 TAAT_ELECO
42 6 5.3 127 1 P044_RAT
43 6 5.3 129 1 CARP_RAT
44 6 5.3 130 1 KR3A_SHEEP
45 6 5.3 131 1 KRA3_SHEEP

ALIGNMENTS

RESULT 1
NIGM_BOVIN
ID NIGM_BOVIN STANDARD; PRT; 108 AA.
AC Q02374;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE AGG SUBUNIT PRECURSOR (EC 1.6.5.3)
DE (EC 1.6.99.3) (COMPLEX I-AGGG) (CI-AGGG).
GN NDUF82.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 37-65.
RC TISSUE=Heart;
RA MEDLINE=92389317; PubMed=1518044;
RX Walker J.E., Arizmendi J.M., Dupuis A., Fearnley I.M., Finel M.,
RT Medd S.M., Pilkington S.J., Runswick M.J., Skehel J.M.;
RT "Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from
RT bovine heart mitochondria. Application of a novel strategy for
RT sequencing proteins using the polymerase chain reaction.";
RL J. Mol. Biol. 226:1051-1072(1992).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC -!- CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X63216; CAA44901.1; -;
CC PIR; S28241; S28241.
DR Oxidoreductase; NAD; 36
KW TRANSIT 1
FT CHAIN 37 108
FT SUBUNIT
SQ SEQUENCE 108 AA; 12282 MW; 0F0AD8B6A38120C7 CRC64;

Query Match 7.1%; Score 8; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 ARAAGARG 15
| | | | | | | |
Db 26 ARAAGARG 33

```

RESULT 2
PSPN_HUMAN
ID PSPN_HUMAN STANDARD; PRT; 156 AA.
AC O60542;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PERSEPHIN PRECURSOR (PSP).
GN PSPN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=98150950; PubMed=9491986;
RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,
RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,
RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,
RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,
RA Klein R.D., Poulsen K., Gray C., Garcés A., Henderson C.E.,
RA Phillips H.S., Johnson E.M.;
RT "Persephin, a novel neurotrophic factor related to GDNF and
RT neurturin.";
RL Neuron 20:245-253(1998).
CC -1- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC
CC DOPAMINERGIC AND MOTOR NEURONS.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; AF040962; AAC39640.1; -
CC HSSP; Q07731; LAGO.
CC MIM; 602921; -.
CC
CC InterPro: IPR001839; TGF-beta.
CC SMART; SM00204; TGFB; 1.
CC PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
KW Growth factor; Signal.
FT SIGNAL 1 21
FT CHAIN 22 156
FT DISULFID 66 124 BY SIMILARITY.
FT DISULFID 93 152 BY SIMILARITY.
FT DISULFID 97 154 BY SIMILARITY.
FT DISULFID 123 133 INTERCHAIN (BY SIMILARITY).
SQ SEQUENCE 156 AA; 16600 MW; 6547751653A7044A CRC64;

Query Match 7.1%; Score 8; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 PCCRPTRY 86
Db 122 PCCRPTRY 129
IIIIIIII
IIIIIIII

RESULT 3
E2F1_MOUSE
ID E2F1_MOUSE STANDARD; PRT; 430 AA.
AC Q61501;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

```

DE TRANSCRIPTION FACTOR E2F1 (E2F-1).
GN E2F1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RX MEDLINE=98025478; PubMed=9376316;
RA Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L.,
RA Phillips R.A.;
RT "Expression patterns of the E2F family of transcription factors during
RT mouse nervous system development.";
RL Mech. Dev. 66:13-25(1997).
RN [3]
RX FUNCTION IN APOPTOSIS.
RX MEDLINE=98337455; PubMed=9674698;
RA Holmberg C., Helin K., Sehested M., Karlstrom O.;
RA "E2F-1-induced p53-independent apoptosis in transgenic mice.";
RL Oncogene 17:143-155(1998).
CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY
CC WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GCCG.
CC FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS
CC ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE
CC DRTF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE
CC PROGRESSION FROM G1 TO S PHASE. E2F-1 BINDS PREFERENTIALLY RB1
CC PROTEIN IN A CELL-CYCLE DEPENDENT MANNER. IT CAN MEDATE BOTH
CC CELL PROLIFERATION AND P53-DEPENDENT APOPTOSIS.
CC -1- SUBUNIT: COMPONENT OF THE DRTF1/E2F TRANSCRIPTION FACTOR COMPLEX.
CC SPECIFICALLY HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1. DURING
CC THE CELL CYCLE, RB1 BECOMES PHOSPHORYLATED IN MID-TO-LATE G1
CC PHASE. DETACHES FROM THE DRTF1/E2F COMPLEX, RENDERING E2F
CC TRANSCRIPTIONALLY ACTIVE. VIRAL ONCOPROTEINS, NOTABLY E1A, T-
CC ANTIGEN AND HPV E7, ARE CAPABLE OF SEQUESTERING RB PROTEIN, THUS
CC RELEASING THE ACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DEVELOPMENTAL STAGE: IN THE DEVELOPING NERVOUS SYSTEM, FIRST
CC DETECTED IN THE NEURAL TUBE AT DAY 9.5 DPC. BY DAY 10.5, LEVELS
CC INCREASE THROUGHOUT THE BRAIN, WITH HIGHEST LEVELS IN THE
CC HINDBRAIN AND IN THE SPINAL CORD, EXPRESSED ONLY IN THE ROSTRAL
CC HALF. BY DAY 11.5, EXPRESSION FOUND THROUGHOUT THE BRAIN AND
CC SPINAL CORD. FROM DAY 12.5, EXPRESSION RESTRICTED TO THE
CC VENTRICULAR REGIONS OF THE BRAIN, PEAKS AT DAY 13.5 AND DECLINES
CC THEREAFTER. ONLY WEAK EXPRESSION IN THE DEVELOPING SPINAL CORD
CC FROM DAY 11.5-16.5. IN THE DEVELOPING RETINA, EXPRESSION IS
CC CONFINED TO THE UNDIFFERENTIATED RETINOBLASTIC CELL LAYER. IN
CC OTHER DEVELOPING TISSUES, E2F-1 IS EXPRESSED IN KIDNEY, LUNG,
CC LIVER HEPATOCYTES, HEART AND THYMUS. HIGHEST LEVELS IN LIVER.
CC ABSENT IN CHOROID PLEXUS.
CC -1- PTM: PHOSPHORYLATED BY CDK2 AND CYCLIN A-CDK2 IN THE S-PHASE (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; L21973; AAA83217.1; -
CC MGD; MGI:101941; E2f1.
CC InterPro; IPR003316; E2F_TDP.
DR

```

DR Pfam; PF02319; E2F_TDP; 1.
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein;
 KW Phosphorylation; Cell cycle; Apoptosis.
 FT DOMAIN 62 103
 FT DNA_BIND 105 189
 FT DOMAIN 148 169
 FT DOMAIN 153 189
 FT DOMAIN 190 279
 FT DOMAIN 361 430
 FT DOMAIN 402 419
 SQ SEQUENCE 430 AA; 46323 MW; C5DF18AD3B4DFEFA CRC64;

Query Match 7.1%; Score 8; DB 1; Length 430;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LLGAGALR 68
 DB 17 LLGAGALR 24

RESULT 4
 E2F1_HUMAN STANDARD; PRT; 437 AA.
 AC Q01094; 092768; Q13143;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DE TRANSCRIPTION FACTOR E2F1 (E2F-1) (RETINOBLASTOMA BINDING PROTEIN 3)
 DE (RBBP-3) (PRB-BINDING PROTEIN E2F-1) (PBR3) (RETINOBLASTOMA-ASSOCIATED
 DE PROTEIN 1) (RBPAP-1).
 DE E2F1 OR RBBP3.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92346720; PubMed-1638634;
 RA Helin K., Lees J.A., Vidal M., Dyson N.J., Harlow E., Fattaey A.;
 RT "A cDNA encoding a PRB-binding protein with properties of the
 RT transcription factor E2F.";
 RL Cell 70:337-350(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92346721; PubMed-1638635;
 RA Kaelin W.G. Jr., Krek W., Sellers W.R., Decaprio J.A., Ajchenbaum F.,
 RA Fuchs C.S., Chittenden T., Li Y., Farnham P.J., Blumberg P.M.,
 RA Livingston D.M., Fennelton E.K.;
 RT "Expression cloning of a cDNA encoding a retinoblastoma-binding
 RT protein with E2F-like properties.";
 RL Cell 70:351-364(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93078763; PubMed-1448092;
 RA Shan B., Zhu X., Chen P.L., Durfee T., Yang Y., Sharp D., Lee W.H.;
 RT "Molecular cloning of cellular genes encoding
 RT retinoblastoma-associated proteins: identification of a gene with
 RT properties of the transcription factor E2F.";
 RL Mol. Cell. Biol. 12:5620-5631(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97082961; PubMed-8964493;
 RA Neuman E., Sellers W.R.S., McNeil J.A., Lawrence J.B.,
 RA Kaelin W.G. Jr.;
 RT "Structure and partial genomic sequence of the human E2F1 gene.";
 RL Gene 173:163-169(1996).
 RN [5]
 RP SEQUENCE OF 1-111 FROM N.A.
 RX MEDLINE-95047311; PubMed-7958836;
 RA Johnson D.G., Ohtani K., Nevins J.R.;

RT "Autoregulatory control of E2F1 expression in response to positive
 RL and negative regulators of cell cycle progression.";
 RN Genes Dev. 8:1514-1525(1994).
 RP PHOSPHORYLATION.
 RX MEDLINE-95140412; PubMed-7838523;
 RA Kitagawa M., Higashi H., Suzuki-Takahashi I., Segawa K., Hanks S.K.,
 RA Taya Y., Nishimura S., Okuyama A.;
 RT "Phosphorylation of E2F-1 by cyclin A-cdk2.";
 RL Oncogene 10:229-236(1995).
 RN [7]
 RP CYCLIN A/CDK2 BINDING DOMAIN.
 RX MEDLINE-94306513; PubMed-8033208;
 RA Krek W., Ewen M.E., Shirodkar S., Arany Z., Kaelin W.G. Jr.,
 RA Livingston D.M.;
 RT "Negative regulation of the growth-promoting transcription factor
 RT E2F-1 by a stably bound cyclin A-dependent protein kinase.";
 RL Cell 78:161-172(1994).
 RN [8]
 RP DIFFERENTIAL REGULATION BY CYCLIN/CDK2 KINASES.
 RX MEDLINE-95047332; PubMed-7958856;
 RA Dynlacht B.D., Flores O., Lees J.A., Harlow E.;
 RT "Differential regulation of E2F transactivation by cyclin/cdk2
 RT complexes.";
 RL Genes Dev. 8:1772-1786(1994).
 RN [9]
 RP REGULATION BY CYCLIN-DEPENDENT KINASES.
 RX MEDLINE-97342624; PubMed-9199321;
 RA Dynlacht B.D., Moberg K., Lees J.A., Harlow E., Zhu L.;
 RT "Specific regulation of E2F family members by cyclin-dependent
 RT kinases.";
 RL Mol. Cell. Biol. 17:3867-3875(1997).
 RN [10]
 RP INHIBITION OF DNA-BINDING.
 RX MEDLINE-95059071; PubMed-7969176;
 RA Xu W., Sheppard K.A., Peng C.-Y., Yee A.S., Plwnica-Worms H.;
 RT "Cyclin A/CDK2 binds directly to E2F-1 and inhibits the DNA-binding
 RT activity of E2F-1/DP-1 by phosphorylation.";
 RL Mol. Cell. Biol. 14:8420-8431(1994).
 RN [11]
 RP FUNCTION IN APOPTOSIS.
 RX MEDLINE-94224788; PubMed-8170954;
 RA Wu X., Levine A.J.;
 RT "P53 and E2F-1 cooperate to mediate apoptosis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:3602-3606(1994).
 RN [12]
 RP TRANSACTIVATION INHIBITION
 RX MEDLINE-94019324; PubMed-8413249;
 RA Helin K., Harlow E., Fattaey A.;
 RT "Inhibition of E2F-1 transactivation by direct binding of the
 RT retinoblastoma protein.";
 RL Mol. Cell. Biol. 13:6501-6508(1993).
 CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY
 CC WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GCGC,
 CC FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS
 CC ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE
 CC DRTF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE
 CC PROGRESSION FROM G1 TO S PHASE. E2F-1 BINDS PREFERENTIALLY RBI
 CC PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER. IT CAN MEDIATE BOTH
 CC CELL PROLIFERATION AND P53-DEPENDENT APOPTOSIS.
 CC -1- SUBUNIT: COMPONENT OF THE DRTF1/E2F TRANSCRIPTION FACTOR COMPLEX.
 CC FORMS HETERODIMERS WITH DP FAMILY MEMBERS. THE E2F-1 COMPLEX BINDS
 CC SPECIFICALLY HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RBL. DURING
 CC THE CELL CYCLE, RBL BECOMES PHOSPHORYLATED IN MID-TO-LATE G1
 CC PHASE, DETACHES FROM THE DRTF1/E2F COMPLEX, RENDERING E2F
 CC TRANSCRIPTIONALLY ACTIVE. VIRAL ONCOPROTEINS, NOTABLY E1A, T-
 CC ANTIGEN AND HPV E7, ARE CAPABLE OF SEQUESTERING RB PROTEIN, THUS
 CC RELEASING THE ACTIVE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- PTM: PHOSPHORYLATED BY CDK2 AND CYCLIN A-CDK2 IN THE S-PHASE.
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE BETA-DEFENSIN 1 PRECURSOR (RHBD-1) (DEFENSIN, BETA 1).
GN DEFBI.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Kwok J., Hurlock G., Wu X., Penland C., Wine J.J.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HAS BACTERICIDAL ACTIVITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF014016; AAC66344.1; -.
DR InterPro: IPR001271; Defensin.
DR Pfam: PF00711; Defensin_beta.1.
DR SMART; SM00048; DEFSN; 1.
KW Antibiotic; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 32 BY SIMILARITY.
FT CHAIN 33 68 BETA-DEFENSIN 1.
FT DISULFID 37 66 BY SIMILARITY.
FT DISULFID 44 59 BY SIMILARITY.
FT DISULFID 49 67 BY SIMILARITY.
FT SEQUENCE 68 AA; 7572 MW; 937336DEC3B0435E CRC64;
SQ
Query Match 6.2%; Score 7; DB 1; Length 68;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 GLGHRSD 35
Db 27 GLGHRSD 33
RESULT 7
ID NRTN_HUMAN STANDARD; PRT; 197 AA.
AC Q99748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEURTURIN PRECURSOR.
GN NRTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-97100947; PubMed-8945474;
RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,
RA Crendon D.J., Johnson E.M. Jr., Milbrandt J.;
RT "Neurturin, a relative of glial-cell-line-derived neurotrophic
RT factor".
RL Nature 384:467-470(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Melanoma;
RA Blum H., Bauersachs S., Mewes H.-W., Weill B., Wiemann S.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP VARIANT HSCR SER-96.
RC TISSUE-Peripheral blood lymphocytes;
RA MEDLINE-98367034; PubMed-9700200;
RA Doray B., Salomon R., Amiel J., Pelet A., Touraine R., Billaud M.,
RA Attie T., Bachy B., Munnich A., Lyonnet S.;
RT "Mutation of the RET ligand, neurturin, supports multigenic
RT inheritance in Hirschsprung disease".
RL Hum. Mol. Genet. 7:1449-1452(1998).
CC -!- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.
CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT
CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS
CC HAEMOPOIETIC CELLS.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DISEASE: IN ASSOCIATION WITH MUTATIONS OF RET GENE AND POSSIBLY
CC OTHER LOCI, IT IS INVOLVED IN HIRSCHSPRUNG'S DISEASE (HSCR). THIS
CC GENETIC DISORDER OF NEURAL CREST DEVELOPMENT IS CHARACTERIZED BY
CC THE ABSENCE OF INTRAMURAL GANGLION CELLS IN THE HINDGUT; OFTEN
CC RESULTING IN INTESTINAL OBSTRUCTION.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U78110; AAC50898.1; -.
DR EMBL; AL161995; CAB82327.1; -.
DR HSSP; Q07731; IAGQ.
DR MIM; 602018; -.
DR MIM; 142623; -.
DR InterPro: IPR002400; GF_cysknot.
DR InterPro: IPR001839; TGF-beta.
DR PRINTS; PR00438; GFCYSKNOT.
DR SMART; SM00204; TGF-beta.1.
DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
KW Growth factor; Signal; Disease mutation: Hirschsprung disease.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 95 BY SIMILARITY.
FT CHAIN 96 197 NEURTURIN.
FT DISULFID 103 165 BY SIMILARITY.
FT DISULFID 130 194 BY SIMILARITY.
FT DISULFID 134 196 BY SIMILARITY.
FT DISULFID 164 164 INTERCHAIN (BY SIMILARITY).
FT VARIANT 96 A -> S (IN HSCR; ASSOCIATED TO A RET
FT MUTATION: INCOMPLETE PENETRANCE).
FT /FTID=VAR_009498.
SQ SEQUENCE 197 AA; 22405 MW; 91AFA8C3F8971FD CRC64;
Query Match 6.2%; Score 7; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 78 QPCCRPT 84
Db 162 QPCCRPT 168
RESULT 8
ID UCRL_CHRVI STANDARD; PRT; 207 AA.
AC O31214;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UBIOQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT (EC 1.10.2.2)
DE (RIESKE IRON-SULFUR PROTEIN) (RISP).

GN PETA.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
OC Allochrochromatium
OX NCBI_TaxID=1049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D / ATCC 17899 / DSM 180;
RA Chen Y.L., Dincturk H.B., Qin H., Knaff D.B.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C = Q +
CC 2 FERROCYTOCHROME C.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL).
CC -1- MISCELLANEOUS: THE RIESKE PROTEIN IS A HIGH POTENTIAL 2FE-2S
CC PROTEIN.
CC -1- SIMILARITY: TO RIESKE PROTEINS FROM OTHER SOURCES (MITOCHONDRIA,
CC BACTERIAL, CHLOROPLAST).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF034104; AAB86973.1; -
CC InterPro; IPR001281; Rieske.1;
CC Pfam; PF00355; Rieske.1;
CC PROSITE; PS00199; RIESKE_1; 1.
CC PROSITE; PS00200; RIESKE_2; 1.
CC Electron transport; Inner membrane; Transmembrane; Iron-sulfur;
KW Oxidoreductase.
FT TRANSMEM 24 44 POTENTIAL.
FT METAL 134 134 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 136 136 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 162 162 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 165 165 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT DISULFID 139 164 BY SIMILARITY.
SQ SEQUENCE 207 AA; 22124 MW; C6886EDF25E981CD CRC64;

Query Match 6.2%; Score 7; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 RARAAGA 13
DB 52 RARAAGA 58
|||||||

RESULT 9
RNC_SYNY3
ID RNC_SYNY3 STANDARD; PRT; 231 AA.
AC P74368;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE RIBONUCLEASE III (EC 3.1.26.3) (RNASE III).
GN RNC OR SLR1646.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,

RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-PHOSPHO-
CC MONOESTER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE RNASE III FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90914; BAA18463.1; -
CC InterPro; IPR001159; DS_RBD.
CC Pfam; PF00035; dsrm; 1.
CC Pfam; PF00636; Ribonuclease_3; 1.
CC SMART; SM00358; DSRM; 1.
CC SMART; SM00535; RIBOC; 1.
CC PROSITE; PS0137; DS_RBD; 1.
CC PROSITE; PS00517; RNASE_3.1; 1.
CC PROSITE; PS0142; RNASE_3.2; 1.
KW Hydrolase; Nuclease; Endonuclease; RNA-binding.
FT DOMAIN 211 227 DRBM.
SQ SEQUENCE 231 AA; 25712 MW; 7638BC35E6B46D60 CRC64;

Query Match 6.2%; Score 7; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 RLRSQV 23
DB 67 RLRSQV 73
|||||||

RESULT 10
WT1_SWIMA
ID WT1_SWIMA STANDARD; PRT; 239 AA.
AC P49953;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE WILMS' TUMOR PROTEIN (FRAGMENT).
GN WT1.
OS Smithopsis macroura (Stripe-faced dunnart).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Dasyuromorphia; Dasyuridae; Sminthopsis.
OX NCBI_TaxID=9302;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96068905; PubMed=7478606;
RA Kent J., Coriat A.M., Sharpe P.T., Hastie N., van Heyningen V.;
RT "The evolution of WT1 sequence and expression pattern in the
RT vertebrates.";
RL Oncogene 11:1781-1792(1995).
CC -1- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCG-3' (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS (BY

CC SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER

CC PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X85732; CAA59737.1; -.

CC HSSP; P08046; 1A1L.

CC InterPro: IPR000976; Wilms tumour.

CC InterPro: IPR000822; Znf-C2H2.

CC Pfam; PF02165; WTI; 2.

CC Pfam; PF00096; zf-C2H2; 4.

CC SMART; SM00355; Znf-C2H2; 4.

CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.

CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.

CC ZINC-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;

CC Transcription regulation; Alternative splicing; Anti-oncogene.

CC NON_TER 1

CC DOMAIN 113 228 ZINC FINGERS.

CC ZN_FING 113 137 C2H2-TYPE.

CC ZN_FING 143 167 C2H2-TYPE.

CC ZN_FING 173 195 C2H2-TYPE.

CC ZN_FING 204 228 C2H2-TYPE.

CC VARSPLIC 198 200 MISSING (IN ISOFORM 2).

CC SEQUENCE 239 AA; 27793 MW; 6707678A7259A624 CRC64;

Query Match 6.2%; Score 7; DB 1; Length 239;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 RSEDLVR 39

DB 217 RSEDLVR 223

RESULT 11

YBLI_STRCI STANDARD; PRT; 242 AA.

ID P33653;

AC 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-FEB-1994 (Rel. 28, Last annotation update)

DE HYPOTHETICAL 26.1 KDA PROTEIN IN BLAB 3'REGION.

OS Streptomyces cacaoi.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1898;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-KCC S0352;

RC MEDLINE=92234939; PubMed=1569015;

RA Urabe H., Ogawara H.;

RT "Nucleotide sequence and transcriptional analysis of

RT activator-regulator proteins for beta-lactamase in Streptomyces

RT cacaoi.;"

RL J. Bacteriol. 174:2834-2842(1992).

CC -1- SIMILARITY: TO E.COLI MAZG AND TO PLASMID PIP1100 ERYTHROMYCIN

CC ESTERASE.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; D00937; BAA00775.1; -.

CC KW Hypothetical protein.

CC SQ SEQUENCE 242 AA; 26051 MW; 286B53C75AF90063 CRC64;

Query Match 6.2%; Score 7; DB 1; Length 242;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SRARAAG 12

DB 186 SRARAAG 192

RESULT 12

Y757_METJA STANDARD; PRT; 260 AA.

ID Y757_METJA

AC Q58167;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE HYPOTHETICAL PROTEIN MJ0757.

GN MJ0757.

OS Methanococcus jannaschii.

OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;

OC Methanococcus.

OX NCBI_TaxID=2190;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;

RC MEDLINE=96337999; PubMed=8688087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,

RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,

RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,

RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,

RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,

RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,

RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;

RT "Complete genome sequence of the methanogenic archaeon, Methanococcus

RT jannaschii.;"

RL Science 273:1058-1073(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U67521; AAB98749.1; -.

CC TIGR; MJ0757;

CC InterPro: IPR003745; DUF166.

CC Pfam; PF02593; DUF166; 1.

CC KW Hypothetical protein.

CC SQ SEQUENCE 260 AA; 30434 MW; F0BBD83A4143BF08 CRC64;

Query Match 6.2%; Score 7; DB 1; Length 260;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 PHDLSLA 59

DB 123 PHDLSLA 129

RESULT 13

WTI_ALLMI STANDARD; PRT; 288 AA.

ID WTI_ALLMI

AC P50902;

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE WILMS' TUMOR PROTEIN (FRAGMENT).
 GN WT1.
 OS Alligator mississippiensis (American alligator).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
 OX NCBI_TaxID=8496;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96068905; PubMed=7478606;
 RA Kent J., Coriat A.M., Sharpe P.T., Hastie N., van Heyningen V.;
 RT "The evolution of WT1 sequence and expression pattern in the
 vertebrates.";
 RL Oncogene 11:1781-1792(1995).
 CC -1- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
 CC AND BINDS TO THE DNA SEQUENCE 5'-GGCCCCCGC-3'.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X85730; CAA59735.1; .
 DR HSSP; P08046; IAL1.
 DR InterPro; IPR000976; Wilms.tumour.
 DR InterPro; IPR000822; Znf-C2H2.
 DR Pfam; PF02165; WT1; 2.
 DR Pfam; PF00096; Zf-C2H2; 4.
 DR SMART; SM00355; Znf-C2H2; 4.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 4.
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
 KW Transcription regulation; Alternative splicing.
 FT NON_TER 1
 FT ZN_FING 162 186 C2H2-TYPE.
 FT ZN_FING 192 216 C2H2-TYPE.
 FT ZN_FING 222 244 C2H2-TYPE.
 FT ZN_FING 253 277 C2H2-TYPE.
 FT VARSPLIC 245 247 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 288 AA; 33111 MW; 33E26F7DBE7BFDDE CRC64;

 Query Match 6.2%; Score 7; DB 1; Length 288;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 RSEDLVR 39
 DB 266 RSEDLVR 272

 RESULT 14
 BST1_RAT
 ID BST1_RAT STANDARD; PRT; 319 AA.
 AC Q63072;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ADP-RIBOSYL CYCLASE 2 PRECURSOR (EC 3.2.2.5) (CYCLIC ADP-RIBOSE
 DE HYDROLASE 2) (CADPR HYDROLASE 2) (BONE MARROW STROMAL ANTIGEN 1) (BST-
 DE 1).
 GN BST1.
 OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=Pancreatic islets;
 RX MEDLINE=96096547; PubMed=8522202;
 RA Furuya Y., Takasawa S., Yonekura H., Tanaka T., Takahara J.,
 RA Okamoto H.;
 RT "Cloning of a cDNA encoding rat bone marrow stromal cell antigen 1
 (BST-1) from the islets of Langerhans.";
 RL Gene 165:329-330(1995).
 CC -1- FUNCTION: SYNTHESIZES CYCLIC ADP-RIBOSE. MAY BE INVOLVED IN PRE-B-
 CC CELL GROWTH.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
 CC -1- TISSUE SPECIFICITY: PANCREATIC ISLETS, KIDNEY, SPLEEN, HEART,
 CC THYMUS, INTESTINE AND SALIVARY GLAND.
 CC -1- SIMILARITY: BELONGS TO THE ADP-RIBOSYL CYCLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D49555; BAA08710.1; .
 DR HSSP; P29241; ILBE.
 DR InterPro; IPR003193; Rib_hydrolayse.
 DR Pfam; PF02267; Rib_hydrolayse; 1.
 KW Hydrolase; Glycoprotein; GPI-anchor; Signal.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 294 ADP-RIBOSYL CYCLASE 2.
 FT PROPEP 295 319 POTENTIAL.
 FT LIPID 294 294 GPI-ANCHOR (POTENTIAL).
 FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 319 AA; 35131 MW; 46831685DE2B2472 CRC64;

 Query Match 6.2%; Score 7; DB 1; Length 319;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 ARAAGAR 14
 DB 28 ARAAGAR 34

 RESULT 15
 PIP_STRCO
 ID PIP_STRCO STANDARD; PRT; 323 AA.
 AC Q9S2L4;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROBABLE PROLINE IMINOPEPTIDASE (EC 3.4.11.5) (PIP) (PROLYL
 DE AMINOPEPTIDASE) (PAP).
 GN SC7H2.03C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Saunders D.C., Harris D., Bentley S.D., Parkhill J., Barrell B.G.,
 RA Rajandream M.A.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SPECIFICALLY CATALYZES THE REMOVAL OF N-TERMINAL PROLINE



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:44:40 ; Search time 35.16 Seconds
(without alignments)
425.212 Million cell updates/sec

Title: US-09-357-349-3
Perfect score: 113
Sequence: 1 AGGCGRARAGARGCRLRS.....VNSTWRTVDRLSATACGLG 113

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 425026 seqs, 132305027 residues

Word size : 0
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPREMBL16.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_unclassified.*
13: sp_vertebrate.*
14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|--------|--------------------|
| 1 | 113 | 100.0 | 220 | 4 | O96030 | homo sapien |
| 2 | 113 | 100.0 | 237 | 4 | O95441 | homo sapien |
| 3 | 30 | 26.5 | 125 | 11 | O9QZG3 | Q9zgq3 rattus norv |
| 4 | 26 | 23.0 | 224 | 11 | O9ZOL2 | Q9z012 mus musculu |
| 5 | 9 | 8.0 | 297 | 2 | O9RD48 | Q9rd48 streptomyce |
| 6 | 8 | 7.1 | 228 | 2 | O9XSR8 | Q9xsr8 canis famil |
| 7 | 8 | 7.1 | 228 | 2 | O9L129 | Q9l129 streptomyce |
| 8 | 8 | 7.1 | 378 | 2 | O53236 | Q53236 rhodobacter |
| 9 | 7 | 6.2 | 13 | 4 | O9HAQ9 | Q9haq9 homo sapien |
| 10 | 7 | 6.2 | 90 | 9 | O9XLT7 | Q9xlt7 bacterioph |
| 11 | 7 | 6.2 | 98 | 6 | O9MVT1 | Q9mvt1 oryctolagus |
| 12 | 7 | 6.2 | 121 | 4 | O9H5Y2 | Q9h5y2 homo sapien |
| 13 | 7 | 6.2 | 124 | 13 | O9DES6 | Q9des6 fugu rubrip |
| 14 | 7 | 6.2 | 133 | 2 | O86339 | O86339 mycobacteri |
| 15 | 7 | 6.2 | 133 | 11 | O93866 | Q93866 rattus norv |
| 16 | 7 | 6.2 | 138 | 14 | O81361 | O81361 hepatitis c |
| 17 | 7 | 6.2 | 148 | 2 | O9X4E4 | Q9x4e4 rhodobacter |
| 18 | 7 | 6.2 | 154 | 9 | O21999 | O21999 bacterioph |
| 19 | 7 | 6.2 | 168 | 4 | O16256 | Q16256 homo sapien |

20 7 6.2 176 2 O9RCU0
21 7 6.2 196 2 P71510
22 7 6.2 201 2 Q07640
23 7 6.2 210 1 Q9HRE1
24 7 6.2 213 2 Q9JQY7
25 7 6.2 220 2 Q9KXN4
26 7 6.2 228 3 Q08747
27 7 6.2 232 5 Q9TVF8
28 7 6.2 234 1 Q9YDR0
29 7 6.2 239 2 Q9S440
30 7 6.2 243 2 Q9S2Y8
31 7 6.2 250 5 Q9TVQ4
32 7 6.2 270 4 Q9H738
33 7 6.2 270 5 Q24217
34 7 6.2 275 4 Q9NQY5
35 7 6.2 279 2 Q9RD19
36 7 6.2 284 5 Q9TW49
37 7 6.2 288 11 Q9ESX3
38 7 6.2 303 2 Q9ZBH0
39 7 6.2 306 2 Q9RTL3
40 7 6.2 308 1 Q9YA95
41 7 6.2 310 2 Q9I6S4
42 7 6.2 316 5 Q9TVY0
43 7 6.2 326 10 Q9EHD6
44 7 6.2 340 2 Q9RJ12
45 7 6.2 344 2 O86662

ALIGNMENTS

RESULT 1
O96030
ID O96030 PRELIMINARY; PRT; 220 AA.
AC O96030;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ARTEMIN.
OS ARTN OR EVN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95098192; PubMed=9883723;
RA Balch R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,
RA Simburger K.S., Leltner M.L., Araki T., Johnson E.M. Jr.,
RA Milbrandt J.;
RT "Artemin, a novel member of the GDNF ligand family, supports
RT peripheral and central neurons and signals through the GFRalpha3-RET
RT receptor complex.";
RL Neuron 21:1291-1302(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=BRAIN;
RA Hansen C., Blom N., Johansen T.E.;
RT "Neublastin a novel member of the GDNF ligand family.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20050601; PubMed=10583383;
RA Masure S., Geerts H., Cik M., Hoefnagel E., Van den Kieboom G.,
RA Tuytelaars A., Harris S., Lesage A.S., Laysen J.E., van der Helm L.,
RA Verhaaselt P., Von J., Gordon R.D.;
RT "Enovin, a member of the glial cell-line-derived neurotrophic factor
RT (GDNF) family with growth promoting activity on neuronal cells.
RT Existence and tissue-specific expression of different splice
RT variants.";
RL Eur. J. Biochem. 266:892-902(1999).
DR EMBL; AF115765; AAC13109.1; -
DR EMBL; AF109401; AAC98690.1; -

DR EMBL; AFI120274; AAD21075.1; -;
 DR EMBL; AJ245628; CAB52396.1; -;
 DR HSP; O07731; IAGO.
 DR InterPro; IPR001839; -;
 DR SMART; SM00204; TGFb; 1.
 KW Signal.
 SQ SEQUENCE 220 AA; 22906 MW; C47754B19AADCFBB CRC64;

Query Match 100.0%; Score 113; DB 4; Length 220;
 Best Local Similarity 100.0%; Pred. No. 4.7e-99;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAAGARGCRLSQLVPVRLGLGHRSDLVRFRCSCGRRARSPHDLAS 60
 Db 108 AGPGSRAAAGARGCRLSQLVPVRLGLGHRSDLVRFRCSCGRRARSPHDLAS 167

QY 61 LLGAGALRPPGSRPVSPQCCRPTRYEA VSFMDVNSTWRTVDRLSATACGCLG 113
 Db 168 LLGAGALRPPGSRPVSPQCCRPTRYEA VSFMDVNSTWRTVDRLSATACGCLG 220

RESULT 2
 O95441 PRELIMINARY; PRT; 237 AA.
 AC O95441;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAY-2001 (TReMBLrel. 16, Last annotation update)
 DE ARTEMIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99098192; PubMed=9883723;
 RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,
 RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,
 RA Milbrandt J.;
 RT "Artemin, a novel member of the GDNF ligand family, supports
 RT peripheral and central neurons and signals through the GFRalpha3-RET
 RT receptor complex."
 RL Neuron 21:1291-1302(1998).
 DR EMBL; AFI15765; AAD13110.1; -;
 DR HSP; O07731; IAGO.
 DR InterPro; IPR001839; -;
 DR SMART; SM00204; TGFb; 1.
 SQ SEQUENCE 237 AA; 24471 MW; 11C64C4B510CE3AB CRC64;

Query Match 100.0%; Score 113; DB 4; Length 237;
 Best Local Similarity 100.0%; Pred. No. 5e-99;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRAAAGARGCRLSQLVPVRLGLGHRSDLVRFRCSCGRRARSPHDLAS 60
 Db 125 AGPGSRAAAGARGCRLSQLVPVRLGLGHRSDLVRFRCSCGRRARSPHDLAS 184

QY 61 LLGAGALRPPGSRPVSPQCCRPTRYEA VSFMDVNSTWRTVDRLSATACGCLG 113
 Db 185 LLGAGALRPPGSRPVSPQCCRPTRYEA VSFMDVNSTWRTVDRLSATACGCLG 237

RESULT 3
 Q9Q2G3 PRELIMINARY; PRT; 125 AA.
 AC Q9Q2G3;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ARTEMIN (FRAGMENT).
 GN ARTN.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=COCHLEA, SUBSTANTIA NIGRA;
 RA Stover T., Gong T.-W., Cho Y., Altschuler R.A., Lomax M.I.;
 RT "Expression of neurturin, artemin, persephin and their receptors GFRa-
 RT 2 and GFRa-3 in the mature rat cochlea."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AFI84919; AAF01241.1; -;
 DR HSP; Q07731; IAGO.
 FT NON_TER 1
 FT NON_TER 125
 SQ SEQUENCE 125 AA; 12983 MW; 8EDE626E44B83231 CRC64;

Query Match 26.5%; Score 30; DB 11; Length 125;
 Best Local Similarity 100.0%; Pred. No. 9.1e-21;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RFRFCGSCRRARSPHDLASLLGAGALR 68
 Db 83 RFRFCGSCRRARSPHDLASLLGAGALR 112

RESULT 4
 Q920L2 PRELIMINARY; PRT; 224 AA.
 ID Q920L2
 AC Q920L2;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAY-2001 (TReMBLrel. 16, Last annotation update)
 DE NEUROTROPHIC FACTOR ARTEMIN PRECURSOR.
 GN ARTN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99098192; PubMed=9883723;
 RA Baloh R.H., Tansey M.G., Lampe P.A., Fahrner T.J., Enomoto H.,
 RA Simburger K.S., Leitner M.L., Araki T., Johnson E.M. Jr.,
 RA Milbrandt J.;
 RT "Artemin, a novel member of the GDNF ligand family, supports
 RT peripheral and central neurons and signals through the GFRalpha3-RET
 RT receptor complex."
 RL Neuron 21:1291-1302(1998).
 DR EMBL; AFI09402; AAC98691.1; -;
 DR HSP; Q07731; IAGO.
 DR MGD; MGI:1333791; Artn.
 DR InterPro; IPR001839; -;
 DR SMART; SM00204; TGFb; 1.
 FT CHAIN 112 224
 SQ SEQUENCE 224 AA; 23726 MW; 3328FB794581DF0B CRC64;

Query Match 23.0%; Score 26; DB 11; Length 224;
 Best Local Similarity 100.0%; Pred. No. 8.8e-17;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 SOPCCRTRYEA VSFMDVNSTWRTVD 102
 Db 188 SOPCCRTRYEA VSFMDVNSTWRTVD 213

RESULT 5
 Q9RD48 PRELIMINARY; PRT; 297 AA.
 ID Q9RD48
 AC Q9RD48;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE HYPOTHETICAL 31.2 KDA PROTEIN.
 GN SCF56.19.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Murphy L., Harris D.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL133424; CAB62764.1; -;
 DR InterPro; IPR000379; -;
 DR InterPro; IPR000734; -;
 DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 297 AA; 31166 MW; 412FF60F1F9428D1 CRC64;

Query Match 8.0%; Score 9; DB 2; Length 297;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 SRARAAGAR 14
 |||||
 DB 215 SRARAAGAR 223

RESULT 6
 OXSR8 PRELIMINARY; PRT; 70 AA.
 AC OXSR8;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DE HYPOTHETICAL 7.0 KDA PROTEIN (FRAGMENT).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=THYROID;
 RA Pichon B., Mercan D., Pouillon V., Christophe-Hobertus C.,
 RA Christophe D.;
 RT "Functional cloning of nuclear proteins and nuclear targeting
 sequences.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ388550; CAB46849.1; -;
 KW Hypothetical protein.
 FT NON-TER 1
 FT 70
 SQ SEQUENCE 70 AA; 7049 MW; 3E59AF2DE4FCEFC9 CRC64;

Query Match 7.1%; Score 8; DB 6; Length 70;
 Best Local Similarity 100.0%; Pred. No. 3.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 4 PGSRARAA 11
 |||||
 DB 7 PGSRARAA 14
 RESULT 7
 OY 09L129 PRELIMINARY; PRT; 228 AA.
 AC OY 09L129;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 23.0 KDA PROTEIN.
 GN SC6D11.26.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL158061; CAB76349.1; -;
 DR InterPro; IPR002808; -;
 DR Pfam; PF01955; DUF105; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 228 AA; 22998 MW; 74FCE4F9899095FA CRC64;

Query Match 7.1%; Score 8; DB 2; Length 228;
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 ARAAGARG 15
 |||||
 DB 81 ARAAGARG 88

RESULT 8
 OY 053236 PRELIMINARY; PRT; 378 AA.
 AC OY 053236;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE TORF PROTEIN.
 GN TORF.
 OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Rhodobacter.
 OX NCBI_TaxID=1063;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WS8;
 RA Goodfellow I.G., Woolley K.J., Sockett R.E.S.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS A SIGMA-54 FACTOR INTERACTION ATP-BINDING
DOMAIN.

DR EMBL: X98694; CAA67253.1; -;
DR InterPro: IPR002078; -;
DR InterPro: IPR003593; -;
DR Pfam: PF00158; sigma54; 1.
DR PROSITE: PS00676; SIGMA54_INTERACT_2; 1.
DR PROSITE: PS00688; SIGMA54_INTERACT_3; 1.
DR PROSITE: PS0045; SIGMA54_INTERACT_4; 1.
DR SMART: SM00382; AAA; 1.
KW ATP-binding; DNA-binding; Transcription regulation.
SQ SEQUENCE 378 AA; 39955 MW; 4FC4B742226DDA84 CRC64;

Query Match 7.1%; Score 8; DB 2; Length 378;

Best Local Similarity 100.0%; Pred. No. 14; Mismatches 0; Indels 0; Gaps 0;

QY 8 ARAAGARG 15
Db 59 ARAAGARG 66
|||||

RESULT 9

Q9HAQ9 PRELIMINARY; PRT; 13 AA.
AC Q9HAQ9; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PROSAPOSIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hukova H., Cervenkova M., Ledvinova J., Poupetova H., Elleder,
RA Hrebicek M., Paton B., Tochackova M., Befekadu A., Berna L.,
RA Harzer K.;
RT "Prosaposin deficiency due to 1bp deletion in the saposin B domain.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF307850; AAG31635.1; -;
FT NON_TER 13
SQ SEQUENCE 13 AA; 1383 MW; 57585B6C1784D727 CRC64;

Query Match 6.2%; Score 7; DB 4; Length 13;

Best Local Similarity 100.0%; Pred. No. 8.3; Mismatches 0; Indels 0; Gaps 0;

QY 58 LASLLGA 64
Db 7 LASLLGA 13
|||||

RESULT 10

Q9ZXL7 PRELIMINARY; PRT; 90 AA.
AC Q9ZXL7; 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE ORE10.
OS bacteriophage phi CTX.
OC Viruses.
OX NCBI_TaxID=35343;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PHICTX-C;
RA Nakayama K., Hayashi T.;
RT "Whole genome sequence of Pseudomonas aeruginosa cytotoxin-converting
phage; phiCTX.";

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-PHICTX-C;
RX MEDLINE=90014160; PubMed=2507866;
RA Hayashi T., Kamio Y., Hishinuma F., Usami Y., Titani K., Terawaki Y.;
RT "Pseudomonas aeruginosa cytotoxin: the nucleotide sequence of the gene
and the mechanism of activation of the protoxin.";
RL Mol. Microbiol. 3:861-868(1989).
DR EMBL: AB008550; BAA36236.1; -;
SQ SEQUENCE 90 AA; 9698 MW; 7E8C73C84E24F4D4 CRC64;

Query Match 6.2%; Score 7; DB 9; Length 90;

Best Local Similarity 100.0%; Pred. No. 39; Mismatches 0; Indels 0; Gaps 0;

QY 58 LASLLGA 64
Db 38 LASLLGA 44
|||||

RESULT 11

Q9MYT1 PRELIMINARY; PRT; 98 AA.
AC Q9MYT1; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE NUCLEAR FACTOR OF ACTIVATED T-CELLS (FRAGMENT).
GN NFATX.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=HEART;
RA Sayeed R.A., Grace A.A., Vandenberg J.I.;
RT "Immediate-early gene response to acute pressure-overload in the
rabbit heart.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ291311; CAB94843.1; -;
FT NON_TER 1
FT NON_TER 98
SQ SEQUENCE 98 AA; 10519 MW; BB499E5E307018C5 CRC64;

Query Match 6.2%; Score 7; DB 6; Length 98;

Best Local Similarity 100.0%; Pred. No. 42; Mismatches 0; Indels 0; Gaps 0;

QY 69 PPGSRP 75
Db 15 PPGSRP 21
|||||

RESULT 12

Q9H5Y2 PRELIMINARY; PRT; 121 AA.
AC Q9H5Y2; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CDNA: FLJ22792 F1S, CLONE KAI2274.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ILEAL MUCOSA;
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,

RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
 RA Yanada K., Fujii Y., Ozaki K., Hirao M., Omori Y., Ota T., Suzuki Y.,
 RA Odayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
 RA Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK026445; BAB15485.1; -;
 SQ SEQUENCE 121 AA; 14290 MW; F9E41C5D2D05E114 CRC64;

Query Match 6.2%; Score 7; DB 4; Length 121;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 CRLRSOL 22
 Db 8 CRLRSOL 14
 |||||

RESULT 13
 Q9DES6 PRELIMINARY; PRT; 124 AA.
 AC Q9DES6;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE P13
 GN CDKN2A/B.

OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Takifugu.
 OX NCBI_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gilley J.N., Fried M.;
 RT "No ARF and only a single CDKN2A-like gene at the Fugu locus homologous
 to the mammalian CDKN2A(p16) tumour suppressor/CDKN2B(p15) region.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ250231; CAC12808.1; -;
 SQ SEQUENCE 124 AA; 13073 MW; 95C78DF9C366EE62 CRC64;

Query Match 6.2%; Score 7; DB 13; Length 124;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 SLLGAGA 66
 Db 21 SLLGAGA 27
 |||||

RESULT 14
 O86339 PRELIMINARY; PRT; 133 AA.
 AC O86339;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE HYPOTHETICAL 13.7 KDA PROTEIN.
 GN RV2060 OR MTFV019.01.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA Parkhill J.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z97984; CAB10786.1; -;
 DR Tuberculist; RV2060; -;
 KW Hypothetical protein.
 SQ SEQUENCE 133 AA; 13665 MW; 6629F5903B64B6E6 CRC64;

Query Match 6.2%; Score 7; DB 2; Length 133;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 VPVRALG 29
 Db 35 VPVRALG 41
 |||||

RESULT 15
 Q63866 PRELIMINARY; PRT; 133 AA.
 AC Q63866;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE WTL ZINC-FINGER HOMOLOG (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93304850; PubMed=8317848;
 RA Madden S.L., Rauscher F.J. III;
 RT "Positive and negative regulation of transcription and cell growth
 mediated by the EGR family of zinc-finger gene products.";
 RL Ann. N. Y. Acad. Sci. 684:75-84(1993).
 DR EMBL; S63358; AAB27319.1; -;
 DR HSSP; P08046; IAAV
 DR InterPro; IPR000822; -;
 DR Pfam; PF00096; zf-C2H2; 4.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
 DR SMART; SM00355; Znf_C2H2; 1.
 KW DNA-binding; Metal-binding; Zinc-finger.
 FT NON_TER 1
 SQ SEQUENCE 133 AA; 16132 MW; 1EDB6DF3C02DB5D0 CRC64;

Query Match 6.2%; Score 7; DB 11; Length 133;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 RSEDLVR 39
 Db 111 RSEDLVR 117
 |||||

Search completed: August 16, 2001, 15:49:24
 Job time: 284 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:39:11 ; Search time 33.13 Seconds
(without alignments)
254.354 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 747

Sequence: 1 PPOSPRAPPAPPAPPALPR.....VNSTWRTVDRSLATACGCLG 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0601.*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 747 | 100.0 | 139 | 21 | Partial human enov |
| 2 | 747 | 100.0 | 140 | 21 | A third predicted |
| 3 | 747 | 100.0 | 140 | 21 | Amino acid sequenc |
| 4 | 747 | 100.0 | 159 | 21 | Protein-2 comprisi |
| 5 | 747 | 100.0 | 220 | 21 | Amino acid sequenc |
| 6 | 747 | 100.0 | 220 | 21 | Short splice varia |
| 7 | 747 | 100.0 | 220 | 21 | A human pre-pro-ne |
| 8 | 747 | 100.0 | 220 | 22 | Human PPO3562 prot |
| 9 | 747 | 100.0 | 228 | 21 | A human GDNF-relat |
| 10 | 747 | 100.0 | 228 | 21 | Long splice varian |
| 11 | 747 | 100.0 | 229 | 21 | Amino acid sequenc |

| | | | | | | |
|----|-------|-------|-----|----|-----------|--------------------|
| 12 | 747 | 100.0 | 237 | 21 | AAV84585 | Alternatively spli |
| 13 | 747 | 100.0 | 237 | 21 | AAV92037 | Human artemin (GDN |
| 14 | 747 | 100.0 | 536 | 21 | AAV84592 | Amino acid sequenc |
| 15 | 735 | 98.4 | 140 | 21 | AAV68707 | Amino acid sequenc |
| 16 | 734 | 98.3 | 237 | 21 | AAV68706 | A human neublastin |
| 17 | 615 | 82.3 | 224 | 21 | AAV93558 | A murine GDNF-rela |
| 18 | 615 | 82.3 | 224 | 21 | AAV84584 | A murine neublasti |
| 19 | 615 | 82.3 | 224 | 21 | AAV68714 | A murine neublasti |
| 20 | 614 | 82.2 | 116 | 21 | AAV84587 | A second predicted |
| 21 | 614 | 82.2 | 116 | 21 | AAV68712 | Amino acid sequenc |
| 22 | 602 | 80.6 | 116 | 21 | AAV68708 | Amino acid sequenc |
| 23 | 601 | 80.5 | 113 | 21 | AAV84586 | A first predicted |
| 24 | 601 | 80.5 | 113 | 21 | AAV68713 | Amino acid sequenc |
| 25 | 600.5 | 80.4 | 222 | 21 | AAV93560 | Consensus GDNF-rel |
| 26 | 589 | 78.8 | 113 | 21 | AAV68709 | Amino acid sequenc |
| 27 | 565.5 | 75.7 | 200 | 21 | AAV68705 | Amino acid sequenc |
| 28 | 537 | 71.9 | 123 | 21 | AAV93565 | Protein encoded by |
| 29 | 515 | 68.9 | 96 | 21 | AAV84589 | Amino acid sequenc |
| 30 | 458 | 61.3 | 160 | 21 | AAV84595 | Amino acid sequenc |
| 31 | 244 | 32.7 | 156 | 20 | AAV16727 | Human pre-pro pers |
| 32 | 244 | 32.7 | 156 | 21 | AAV92038 | Human glial cell d |
| 33 | 241 | 32.3 | 197 | 18 | AAV13716 | Human pre-pro-neur |
| 34 | 241 | 32.3 | 197 | 20 | AAV16637 | WO9914235 Seq ID N |
| 35 | 241 | 32.3 | 197 | 20 | AAV900863 | Human pre-pro-neur |
| 36 | 241 | 32.3 | 197 | 21 | AAV92014 | Human neurturin. |
| 37 | 232.5 | 31.1 | 156 | 20 | AAV16721 | Murine pre-pro per |
| 38 | 231 | 30.9 | 195 | 18 | AAV13717 | Mouse pre-pro-neur |
| 39 | 231 | 30.9 | 195 | 20 | AAV16638 | WO9914235 Seq ID N |
| 40 | 230 | 30.8 | 185 | 18 | AAV26680 | Mature mouse perse |
| 41 | 230 | 30.8 | 185 | 20 | AAV16692 | WO9914235 Seq ID N |
| 42 | 228.5 | 30.6 | 134 | 18 | AAV30067 | Mouse persephin. |
| 43 | 228.5 | 30.6 | 134 | 20 | AAV16663 | WO9914235 Seq ID N |
| 44 | 228.5 | 30.6 | 142 | 20 | AAV16681 | WO9914235 Seq ID N |
| 45 | 228 | 30.5 | 96 | 18 | AAV30075 | Persephin-neurturi |

ALIGNMENTS

RESULT 1
AAV45011
ID AAV45011 standard; Protein; 139 AA.

XX AAV45011;

XX AC

XX 17-MAY-2000 (first entry)

XX Partial human enovin protein.

XX Enovin; EVN; neurotrophic growth factor; chromosome lp31.3-32;
glial cell-line derived neurotrophic factor; GDNF; neuroprotective;
GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;
antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;
antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;
Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;
endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;
dyspepsia; constipation; intestinal atony; emesis; diarrhoea;
Crohn's disease; bowel hypersensitivity.

XX Homo sapiens.

XX Key Location/Qualifiers
XX Peptide 1..26
XX /label= Pro-sequence
XX /note= "Partial pro sequence of enovin"
XX Protein 27..139
XX /label= Mature_Enovin
XX /note= "Homologous to GDNF, Neurturin and Persephin"
XX Modified-site 121..123
XX /note= "Asn is N-glycosylated"

XX WO200004050-A2.

XX

PD 27-JAN-2000.
 PF 14-JUL-1999; 99WO-EP05031.
 XX
 XX 14-JUL-1998; 98GB-0015283.
 PR 12-FEB-1999; 99US-0248772.
 PR 08-JUN-1999; 99US-0327668.
 XX
 PA (JANC) JANSSEN PHARM NV.
 XX
 PI Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck IAL;
 XX
 DR WPI; 2000-182404/16.
 DR N-PSDB; AAZ50705.
 XX
 XX Novel human neurotrophic growth factor designated enovin used to treat
 PT neurological disorders, neuronal disorders, peripheral neuropathy,
 PT brain injury, nervous system disorders, prion associated and
 PT gastrointestinal diseases -
 XX
 PS Claim 1; Fig 1; 125pp; English.
 XX
 CC The present protein sequence is that of human enovin comprising
 CC partial pro sequence. Enovin (EVN) is a neurotrophic growth
 CC factor, that belongs to glial cell-line derived neurotrophic
 CC factor (GDNF) family. It binds to GDNF family receptor
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and
 CC prostate. It has nootropic, analgesic, neuroprotective, antiarithmetic,
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,
 CC laxative and antiemetic activity. It can be used to treat neurological
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
 XX
 SQ Sequence 139 AA;
 Query Match 100.0%; Score 747; DB 21; Length 139;
 Best Local Similarity 100.0%; Pred. No. 1.7e-57;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPOPSRPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60
 XX
 Db 1 PPOPSRPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60
 QY 61 DELVRFRCGSGCRRRARSPHDLASLLGAGALRPPGSRPVSPQCCRPTRYEAVSFMDV 120
 XX
 Db 61 delvrfrcsgscrrarsphdlsllagallrppgsrpvsgpcrptryeavsfmdv 120
 QY 121 NSTWRTVDRLSATACGCLG 139
 XX
 Db 121 nstwrtdrlsatacgcig 139
 RESULT 2
 AAY84588
 ID AAY84588 standard; Protein; 140 AA.
 XX
 AC AAY84588;
 XX
 DT 25-JUL-2000 (first entry)
 XX
 DE A third predicted human mature artemin polypeptide.
 XX
 KW Human; artemin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;
 KW
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.
 OS Homo sapiens.
 XX
 PN WO200018799-A1.
 XX
 PD 06-APR-2000.
 XX
 XX 29-SEP-1999; 99WO-US22604.
 XX
 PR 29-SEP-1998; 98US-0163283.
 PR 12-NOV-1998; 98US-0108148.
 PR 22-DEC-1998; 98US-0218698.
 XX
 XX (UNIW) UNIV WASHINGTON.
 XX
 PI Milbrandt JD, Baloh RH;
 XX
 DR WPI; 2000-293109/25.
 DR N-PSDB; AAA12546.
 XX
 XX Isolated artemin growth factor proteins and the nucleic acids that
 PT encode them, useful for treating a range of degenerative neuronal
 PT disorders such as Parkinson's disease and Huntington's disease -
 XX
 PS Claim 4; Fig 3C; 96pp; English.
 XX
 CC The present sequence represents a predicted mature human artemin growth
 CC factor protein. Artemin is a neurotrophic factor that belongs to the
 CC GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin
 CC family of growth factors and promotes differentiation, maintains mature
 CC phenotypes and provides trophic support, promoting growth and survival of
 CC neurons. Artemin promotes the survival of trigeminal ganglion neurons,
 CC nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin
 CC is the only member of the GDNF family that binds to GFR-alpha (growth
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (ret protein-
 CC tyrosine kinase) receptor complex and additionally, like GDNF and
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin
 CC polypeptides and polynucleotides are administered to treat peripheral
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.
 CC blastoma), multiple sclerosis, infection or enteric disease (e.g.
 CC idiopathic constipation or constipation associated with Parkinson's
 CC disease, spinal cord injury or use of opiate pain killers). They may
 CC also be used to treat a patient suffering from small cell lung carcinoma.
 XX
 SQ Sequence 140 AA;
 Query Match 100.0%; Score 747; DB 21; Length 140;
 Best Local Similarity 100.0%; Pred. No. 1.7e-57;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPOPSRPPAPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60
 XX
 Db 2 PPOPSRPPAPPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 61
 QY 61 DELVRFRCGSGCRRRARSPHDLASLLGAGALRPPGSRPVSPQCCRPTRYEAVSFMDV 120
 XX
 Db 62 delvrfrcsgscrrarsphdlsllagallrppgsrpvsgpcrptryeavsfmdv 121
 QY 121 NSTWRTVDRLSATACGCLG 139
 XX
 Db 122 nstwrtdrlsatacgcig 140
 RESULT 3
 AAY68711
 ID AAY68711 standard; Protein; 140 AA.
 XX

06-JUL-1998; 98DK-0000904.
 09-JUL-1998; 98US-0092229.
 19-AUG-1998; 98DK-0001048.
 25-AUG-1998; 98US-0097774.
 06-OCT-1998; 98DK-0001265.
 13-OCT-1998; 98US-0103908.
 02-JUL-1999; 99US-0347613.
 (NEUR-) NEUROSEARCH AS.
 JOHNSON TE, Blom N, Hansen C;
 WPI; 2000-171013/15.
 N-PSDB; AAZ60563.
 New isolated polypeptides, used for treating e.g. neurodegenerative disease or disorder, neuronal damage or neuronal disorder of the peripheral nervous system, the medulla or the spinal cord -
 Claim 14; Page 97; 106pp; English.
 The present sequence represents a neurotrophic factor designated neublastin. Neublastin is a member of the glial cell line-derived neurotrophic factor sub-class of the transforming growth factor-beta superfamily of neurotrophic factors. Neublastin exhibits high affinity for the GFR-alpha3-RET receptor complex. The polypeptides can be used for treating a neurodegenerative disease or disorder, cerebral ischemic neuronal damage, traumatic brain injury, peripheral neuropathy, Alzheimer's disease, Huntington's disease, Parkinson's disease, Parkinson-Plus syndromes, progressive Supranuclear Palsy, Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian parkinsonism, dementia complex, amyotrophic lateral sclerosis, memory impairment, or a neuronal disorder of the peripheral nervous system, the medulla or the spinal cord. They can also be used for treating various neuropathies. They can also be used for treating ischemic stroke, acute brain injury, acute spinal cord injury, nervous system tumors, multiple sclerosis, exposure to neurotoxins, metabolic diseases such as diabetes or renal dysfunctions and damage caused by infectious agents, or various disorders in the eye.

Query Match 100.0%; Score 747; DB 21; Length 220;
 Best Local Similarity 100.0%; Pred. No. 2.6e-57;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPQSRPAPPAPPALPRGGRARAGGPGSRARAAGRCRLRSQLVPRVRLGLGHS 60
 Db 82 PPQSRPAPPAPPALPRGGRARAGGPGSRARAAGRCRLRSQLVPRVRLGLGHS 141
 QY 61 DELVRFRCGSCRRARSHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120
 Db 142 delvrfrcgscrrarsphdlsllagallrppgsrpsvpcrrptryeavsfmdv 201
 QY 121 NSTWRTVDRLSATACGCLG 139
 Db 202 nstwtvdrslsatacgclg 220

RESULT 8
 ID AAB50978
 XX AAB50978 standard; Protein: 220 AA.
 AC AAB50978;
 XX
 DT 21-MAR-2001 (first entry)
 XX
 DE Human PRO3562 protein.
 KW Human; PRO: cytostatic; neurotropic; neuroprotective; respiratory general;
 KW antinflammatory; antiangiogenic; immunosuppressive; immunostimulant;
 KW PRO agonist; cancer; inflammatory disorder; immunological disorder.

XX Homo sapiens.
 OS WO200073348-A2.
 PN
 XX 07-DEC-2000.
 XX
 XX 30-MAY-2000; 2000WO-US14941.
 XX
 XX 02-JUN-1999; 99WO-US12252.
 PR 22-JUN-1999; 99US-0140650.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 29-OCT-1999; 99US-0162506.
 PR 30-NOV-1999; 99WO-US28313.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30999.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03365.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 03-MAR-2000; 2000US-0187202.
 PR 10-MAR-2000; 2000WO-US06319.
 PR 13-MAR-2000; 2000WO-US06884.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 XX (GETH) GENENTECH INC.
 PA Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC;
 PI Shelton DL, Smith V, Watanabe CK, Wood WI;
 XX WPI; 2001-016509/02.
 DR N-PSDB; AAC91580.
 XX
 PT Twenty eight nucleic acids encoding PRO polypeptides which are useful for treating various tumors, e.g. breast cancer, and other inflammatory, angiogenic and immunological disorders -
 XX
 PS Claim 31; Fig 56; 188pp; English.
 XX
 CC The present sequence is one of twenty eight novel PRO polypeptides. The PRO polypeptides and their agonists, including antibodies, peptides, and small molecule agonists, may be used to treat various tumors, e.g., cancers such as breast cancer, ovarian cancer, renal cancer, colorectal cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer, central nervous system cancer, melanoma or leukaemia. They are also useful for treating other disorders such as neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal and blastocoeleic disorders, and inflammatory, angiogenic and immunological disorders.
 XX
 SQ Sequence 220 AA;

Query Match 100.0%; Score 747; DB 22; Length 220;
 Best Local Similarity 100.0%; Pred. No. 2.6e-57;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPQSRPAPPAPPALPRGGRARAGGPGSRARAAGRCRLRSQLVPRVRLGLGHS 60
 Db 82 PPQSRPAPPAPPALPRGGRARAGGPGSRARAAGRCRLRSQLVPRVRLGLGHS 141
 QY 61 DELVRFRCGSCRRARSHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120
 Db 142 delvrfrcgscrrarsphdlsllagallrppgsrpsvpcrrptryeavsfmdv 201
 QY 121 NSTWRTVDRLSATACGCLG 139


```
Db 202 nstwtvdrisatcgclg 220
|||||
RESULT 9
ID AAY93559 standard; Protein; 228 AA.
AC AAY93559;
XX
DT 25-SEP-2000 (first entry)
XX
DE A human GDNF-related neurotrophic factor 4 (GRNF4).
XX
KW GDNF; glial cell line-derived neurotrophic factor; GFRalpha-3;
KW GDNF-related neurotrophic factor 4; GRNF4; GDNF family receptor-alpha-3;
KW Parkinson's disease; Alzheimer's disease; anyotrophic lateral sclerosis;
KW incontinence; bone loss; osteoporosis; osteogenesis imperfecta;
KW hypercalcemia; nerve damage; stroke; cancer; dideoxycytidine; AIDS;
KW chronic metabolic disease; renal dysfunction.
XX
OS Homo sapiens.
XX
PN W0200034475-A2.
XX
PD 15-JUN-2000.
XX
PF 08-DEC-1999; 99WO-0528975.
XX
PR 09-DEC-1998; 98US-0111626.
XX
PA (AMGE-) AMGEN INC.
XX
PI Simonet WS, Asuncion EJ, Min H, Jing S;
XX
DR WPI; 2000-423421/36.
XX
DR N-PSDB; AAA46615.
XX
PT New glial cell line-derived neurotrophic factor-related neurotrophic
PT factor 4 useful for treating neurodegenerative disease such as
PT Parkinson's disease and for treating nerve damage caused by physical
PT injury and other metabolic diseases
XX
PS Claim 1; Fig 7; 135pp; English.
XX
CC The present sequence represents a human GDNF (glial cell line-derived
CC neurotrophic factor)-related neurotrophic factor 4 (GRNF4) protein.
CC The GRNF4 polypeptide is capable of binding a GDNF family
CC receptor-alpha-3 (GFRalpha-3). The GRNF4 polynucleotides may be
CC used for in vitro GRNF4 protein production as well as in cell therapy
CC or gene therapy applications. GRNF4 protein product may be used in
CC treating Parkinson's disease, Alzheimer's disease, anyotrophic
CC lateral sclerosis, incontinence, diseases associated with bone loss
CC (e.g. osteoporosis, osteogenesis imperfecta or hypercalcemia of
CC malignancy). GRNF4 protein products may also be used in the treatment
CC of nerve damage which may occur to one or more types of nerve cells by
CC physical injury, which causes the degeneration of the axonal processes
CC and/or nerve cell bodies near the site of injury, temporary or permanent
CC cessation of blood flow to parts of the nervous system, as in stroke,
CC intentional or accidental exposure to neurotoxins, for e.g.
CC chemotherapeutic agents for the treatment of cancer or dideoxycytidine
CC for the treatment of AIDS, chronic metabolic diseases, including
CC diabetes or renal dysfunction.
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 747; DB 21; Length 228;
Best Local Similarity 100.0%; Pred. No. 2.7e-57;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPQSPRAPPAPPSPALPGGRRARAGGPGSRARAGARGCRLRSQLVPRALGLGHS 60
|||||
```


Alternatively spliced human artemin polypeptide.

Human; artemin; growth factor; neurotrophic factor; trophic support; neuron; trigeminal ganglion neuron; nodose ganglion neuron; superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease; peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke; Parkinson's disease; Huntington's disease; acute brain injury; acute spinal cord injury; nervous system tumour; blastoma; multiple sclerosis; infection; enteric disease; idiopathic constipation; Parkinson's disease; small cell lung carcinoma.

Homo sapiens.

OS XX

PN XX

WO200018799-A1.

PD 06-APR-2000.

PF 29-SEP-1999; 99WO-US22604.

XX 29-SEP-1998; 98US-0163283.

PR 12-NOV-1998; 98US-0108148.

PR 22-DEC-1998; 98US-0218698.

XX (UNIW) UNIV WASHINGTON.

PA

XX Milbrandt JD, Baloh RH;

PI

XX WPI; 2000-293109/25.

DR N-PSDB; AAA12542.

XX

XX Isolated artemin growth factor proteins and the nucleic acids that

PT encode them, useful for treating a range of degenerative neuronal

PT disorders such as Parkinson's disease and Huntington's disease -

XX

XX Claim 5; Fig 1D; 96pp; English.

PS

XX The present sequence is encoded by an alternatively spliced human artemin

CC growth factor cDNA. Artemin is a neurotrophic factor that belongs to

CC the GDNF (glial cell line-derived neurotrophic factor)/neurturin/

CC persephin family of growth factors and promotes differentiation,

CC maintains mature phenotype and provides trophic support, promoting

CC growth and survival of neurons. Artemin promotes the survival of

CC trigeminal ganglion neurons, nodose ganglion neurons, superior cervical

CC ganglion neurons and tyrosine-hydroxylase-expressing dopaminergic

CC ventral midbrain neurons. Artemin is the only member of the GDNF family

CC that binds to GFR-alpha (growth factor receptor-alpha) and activates

CC the GFR-alpha3/RET (ret protein-tyrosine kinase) receptor complex and

CC additionally, like GDNF and neurturin, artemin also binds to and

CC activates GFRalpha1/RET. Artemin polypeptides and polynucleotides are

CC administered to treat peripheral neuropathy, amyotrophic lateral

CC sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's

disease, ischemic stroke, acute brain injury, acute spinal cord injury, a nervous system tumour (e.g. blastomas), multiple sclerosis, infection or enteric disease (e.g. idiopathic constipation or constipation associated with Parkinson's disease, spinal cord injury or use of opiate pain killers). They may also be used to treat a patient suffering from small cell lung carcinoma.

Sequence 237 AA;

Query Match 100.0%; Score 747; DB 21; Length 237;

Best Local Similarity 100.0%; Pred. No. 2.8e-57;

Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPQSRAPPPAPPSPALPPGGRNARAGGPGSGRARAGRCRLRSQLVPRVRLGLGHR 60

Db 99 PPQSRAPPPAPPSPALPPGGRNARAGGPGSGRARAGRCRLRSQLVPRVRLGLGHR 158

Qy 61 DELVRFRCSCGSRARRSPHDLASLLGAGALRPVPPGSRPVSPQCCRPTRYEAVSFMDV 120

Db 159 delvrfrcscgrrarsphdlsiasllgagairpppgsrpvsgpcrrtryeavsfmdv 218

Qy 121 NSTWRTVDRLSATACGCLG 139

Db 219 nstwtvdrslsatacgclg 237

RESULT 13

AA92037

ID AAY92037 standard; Protein; 237 AA.

XX

AC AAY92037;

XX

DT 19-JUL-2000 (first entry)

XX

DE Human artemin (GDNF) subunit.

XX

KW human artemin; GDNF; CKGF; mutant; cystine knot growth factor;

KW hairpin loop; infertility.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT Misc-difference 1..143

FT /note= "optionally mutated to increase electrostatic

FT interaction between beta hairpin structure and

FT a receptor"

FT

FT Domain 144..163

FT /label= beta_hairpin_loop_1

FT /note= "mutant optionally comprises one or more

FT substitutions in these residues"

FT

FT Misc-difference 164..208

FT /note= "optionally mutated to increase electrostatic

FT interaction between beta hairpin structure and

FT a receptor"

FT

FT Domain 209..229

FT /label= beta_hairpin_loop_3

FT /note= "mutant optionally comprises one or more

FT substitutions in these residues"

FT

FT Misc-difference 230..237

FT /note= "optionally mutated to increase electrostatic

FT interaction between beta hairpin structure and

FT a receptor"

FT

XX WO200017360-A1.

XX

XX 30-MAR-2000.

XX

XX 19-WAR-1999; 99WO-US05908.

XX

XX 22-SEP-1998; 98WO-US19772.

XX

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX

XX Weintraub BD, Szkludinski MW;

XX

XX WPI; 2000-283585/24.

XX

XX New mutant cystine knot growth factor proteins comprising one or more

XX mutant subunits, useful for treating or preventing diseases e.g.

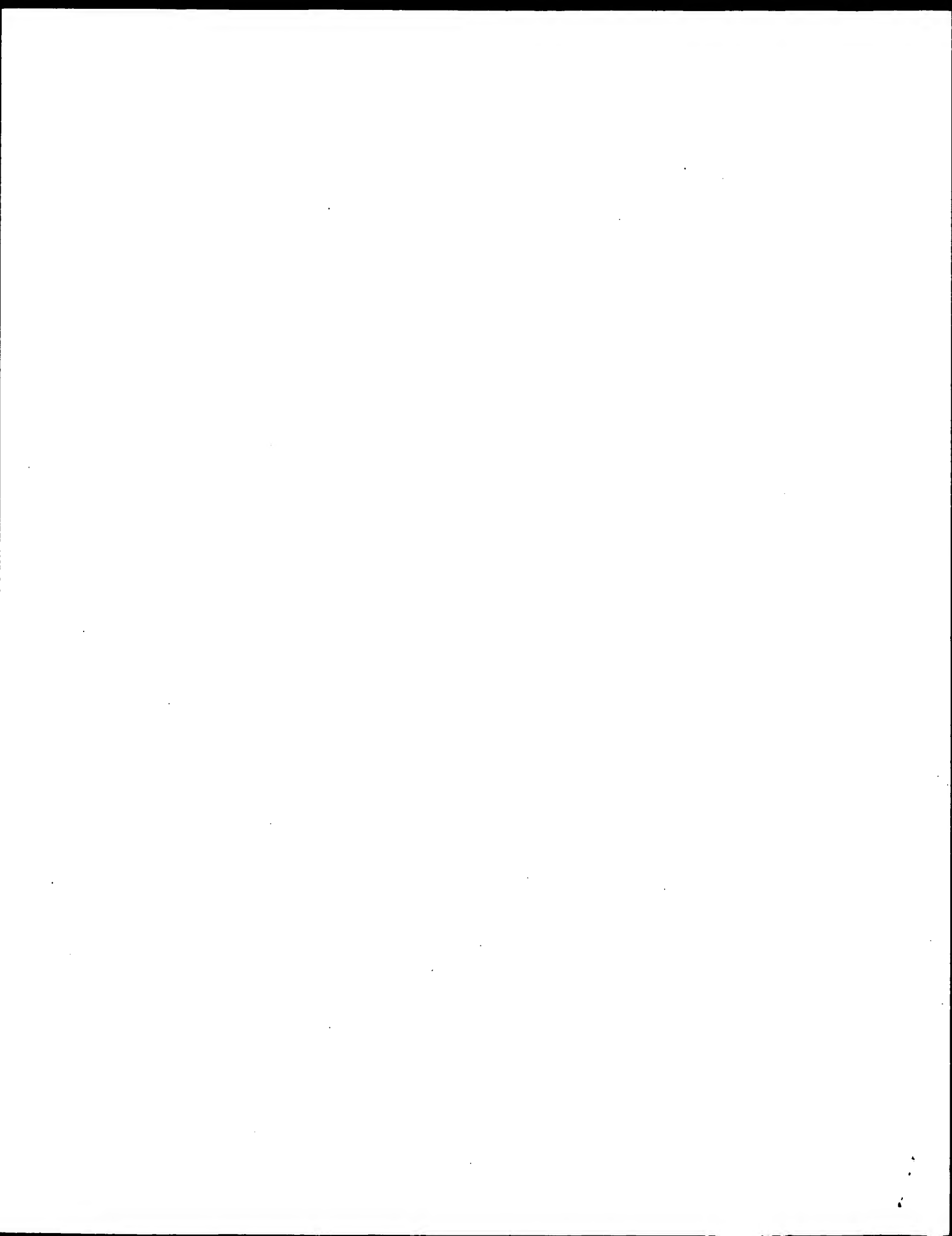
XX hypothyroidism and thyroid cancer

Claim 594; Page 314; 320pp; English.

This is the wild type human artemin (GDNF).

Mutants comprise at least one electrostatic charge altering mutation in a beta hairpin loop, resulting in increased bioactivity.

Mutant cystine knot growth factor (CKGF) proteins comprising one or more mutant subunits and having novel properties or improved pharmacological properties, compared to wild type CKGFs, are claimed. The CKGF superfamily comprises at least four families of growth factors: the glycoprotein hormones, the platelet-derived growth factor (PDGF) family, the neurotrophins and the transforming growth factor-beta family; the families are known to be structurally similar (especially comprising the cystine knot topology) and it was shown that mutations at certain



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:39:37 ; Search time 18.91 Seconds
(without alignments)
151.352 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 747

Sequence: 1 PPQSRPAPPAPPSPALPR.....VNSTWRTVDRLSATACGLG 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
 - 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
 - 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
 - 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
 - 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
 - 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 244 | 32.7 | 133 | 4 | US-08-931-858E-132 |
| 2 | 244 | 32.7 | 156 | 4 | US-08-931-858E-217 |
| 3 | 241 | 32.3 | 144 | 3 | US-08-775-414-81 |
| 4 | 241 | 32.3 | 152 | 3 | US-08-775-414-83 |
| 5 | 241 | 32.3 | 197 | 1 | US-08-519-777-7 |
| 6 | 241 | 32.3 | 197 | 1 | US-08-742-035-7 |
| 7 | 241 | 32.3 | 197 | 2 | US-08-777-019-7 |
| 8 | 241 | 32.3 | 197 | 2 | US-08-777-143-7 |
| 9 | 241 | 32.3 | 197 | 3 | US-09-106-486-5 |
| 10 | 241 | 32.3 | 197 | 3 | US-08-775-414-7 |
| 11 | 241 | 32.3 | 197 | 4 | US-08-931-858E-7 |
| 12 | 241 | 32.3 | 197 | 4 | US-08-981-739-7 |
| 13 | 232.5 | 31.1 | 156 | 4 | US-08-931-858E-185 |
| 14 | 231 | 30.9 | 142 | 3 | US-08-775-414-82 |
| 15 | 231 | 30.9 | 150 | 3 | US-08-775-414-84 |
| 16 | 231 | 30.9 | 195 | 1 | US-08-519-777-8 |
| 17 | 231 | 30.9 | 195 | 1 | US-08-742-035-8 |
| 18 | 231 | 30.9 | 195 | 2 | US-08-777-019-8 |
| 19 | 231 | 30.9 | 195 | 2 | US-08-777-143-8 |
| 20 | 231 | 30.9 | 195 | 3 | US-08-775-414-8 |
| 21 | 231 | 30.9 | 195 | 4 | US-08-931-858E-8 |
| 22 | 231 | 30.9 | 195 | 4 | US-08-981-739-8 |
| 23 | 230 | 30.8 | 185 | 4 | US-08-981-739-133 |
| 24 | 228.5 | 30.6 | 134 | 4 | US-08-981-739-81 |
| 25 | 228.5 | 30.6 | 142 | 4 | US-08-931-858E-111 |
| 26 | 228.5 | 30.6 | 142 | 4 | US-08-981-739-111 |
| 27 | 228 | 30.5 | 96 | 4 | US-08-931-858E-141 |

| | | | | | | |
|----|-------|------|-----|---|--------------------|-------------------|
| 28 | 228 | 30.5 | 96 | 4 | US-08-981-739-141 | Sequence 141, App |
| 29 | 227.5 | 30.5 | 96 | 4 | US-08-931-858E-80 | Sequence 80, Appl |
| 30 | 227.5 | 30.5 | 96 | 4 | US-08-931-858E-187 | Sequence 187, App |
| 31 | 227.5 | 30.5 | 96 | 4 | US-08-981-739-80 | Sequence 80, Appl |
| 32 | 225.5 | 30.2 | 91 | 4 | US-08-931-858E-89 | Sequence 89, Appl |
| 33 | 225.5 | 30.2 | 91 | 4 | US-08-981-739-89 | Sequence 89, Appl |
| 34 | 225.5 | 30.2 | 96 | 4 | US-08-931-858E-221 | Sequence 221, App |
| 35 | 225.5 | 30.2 | 104 | 3 | US-08-775-414-79 | Sequence 79, Appl |
| 36 | 224 | 30.0 | 185 | 4 | US-08-981-739-136 | Sequence 136, App |
| 37 | 223.5 | 29.9 | 89 | 4 | US-08-931-858E-79 | Sequence 79, Appl |
| 38 | 223.5 | 29.9 | 89 | 4 | US-08-931-858E-223 | Sequence 223, App |
| 39 | 223.5 | 29.9 | 89 | 4 | US-08-981-739-79 | Sequence 79, Appl |
| 40 | 222 | 29.7 | 102 | 1 | US-08-519-777-1 | Sequence 1, Appl |
| 41 | 222 | 29.7 | 102 | 1 | US-08-742-035-1 | Sequence 1, Appl |
| 42 | 222 | 29.7 | 102 | 2 | US-08-777-019-1 | Sequence 1, Appl |
| 43 | 222 | 29.7 | 102 | 2 | US-08-777-143-1 | Sequence 1, Appl |
| 44 | 222 | 29.7 | 102 | 3 | US-09-106-486-1 | Sequence 1, Appl |
| 45 | 222 | 29.7 | 102 | 3 | US-08-775-414-1 | Sequence 1, Appl |

ALIGNMENTS

RESULT 1
US-08-931-858E-132
; Sequence 132, Application US/08931858E
; Patent No. 622022
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, EUGENE M
; APPLICANT: MILBRANDT, JEFFREY D
; APPLICANT: KOTZBAUER, PAUL T
; APPLICANT: LAMPE, PATRICIA A
; APPLICANT: KLEIN, ROBERT
; APPLICANT: DESAUVAGE, FRED
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931.858E
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 971486
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-931-858E-132

Query Match 32.7%; Score 244; DB 4; Length 133;
Best Local Similarity 44.8%; Pred. No. 8.9e-16;
Matches 56; Conservative 17; Mismatches 36; Indels 16; Gaps 3;

| | | |
|----|-----|---|
| Qy | 23 | RARAGG-----PGSRARAGACRRLRSQLVPVRALGLHRSDLVRRFCGSGC- R 74 |
| | | : : : : : : : : : : : : : : : : : : : |
| Db | 40 | QVAKAGGTWLTGTHRLARLRRALSGPCQLWSLTLSVAELGLGYASEBKVIIFYCAGSCPR 99 |
| | | : : : : : : : : : : : : : : : : : : : |
| Qy | 75 | RARSPhDLSLASLLGAGALRPPEGSPRVQPCCRPTRYEAVSFMDVNSTWRIVDRLSATA 134 |
| | | : : : : : : : : : : : : : : : : : : : |
| Db | 100 | GARTQHGLALARLGQG-----RAHGGPCCRPTRYTDVAFLDHRRHWRLPOLSAAA 151 |
| Qy | 135 | CGCLG 139 |
| | | : : |

RESULT 4
US-08-775-414-83
; Sequence 83, Application US/08775414
; Patent No. 6090778
; GENERAL INFORMATION:
; APPLICANT: JOHNSON JR., EUGENE M.
; APPLICANT: MILBRANDT, JEFFREY D.
; APPLICANT: KOTZBAUER, PAUL T.
; APPLICANT: LAMPE, PATRICIA A.

;; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
;; NUMBER OF SEQUENCES: 90
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
;; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
;; CITY: ST. LOUIS
;; STATE: MISSOURI
;; COUNTRY: US
;; ZIP: 63105-1817
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/775,414
;; FILING DATE: 31-DEC-1996
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: HOLLAND, DONALD R.
;; REGISTRATION NUMBER: 35,197
;; REFERENCE/DOCKET NUMBER: 965805
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (314) 727-5188
;; TELEFAX: (314) 727-6092
;; INFORMATION FOR SEQ ID NO: 83:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 152 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-775-414-83

Query Match 32.3%; Score 241; DB 3; Length 152;
Best Local Similarity 46.2%; Pred. No. 1.9e-15;
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

Qy 11 PPAPPSALPGGAAAGGPGSARA-AGARGCRLRSQLVVPVRAALGLGHRSDLVRFRC 69
Db 35 PPGP-----RRRAGPRRRRARLARGPCGLRELEVRVSELGLGYASDETVLFRC 85
Qy 70 SGSCRRARSPhdLSLGLGALRPPGSRPVSPQCCRPTRYE-AVSFMDVNSTWRTVD 128
Db 86 AGACEAAARYVDLGLRLRRLR---RRVRAQPCCRPTAYEDEVSLDAHSRYHTVH 142
Qy 129 RLSATACGCL 138
Db 143 ELSARECACV 152

RESULT 5
US-08-519-777-7
; Sequence 7, Application US/08519777
; Patent No. 5739307
; GENERAL INFORMATION:
; APPLICANT: JOHNSON JR., EUGENE M.
; APPLICANT: MILBRANDT, JEFFREY D.
; APPLICANT: KOTZBAUER, PAUL T.
; APPLICANT: LAMPE, PATRICIA A.
; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: US
; ZIP: 63105-1817
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,035
; FILING DATE: 01-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/519,777
; FILING DATE: 28-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 953095
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/519,777
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: HOLLAND, DONALD R.
;; REGISTRATION NUMBER: 35,197
;; REFERENCE/DOCKET NUMBER: 953095
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (314) 727-5188
;; TELEFAX: (314) 727-6092
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 197 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-519-777-7

Query Match 32.3%; Score 241; DB 1; Length 197;
Best Local Similarity 46.2%; Pred. No. 2.6e-15;
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

Qy 11 PPAPPSALPGGAAAGGPGSARA-AGARGCRLRSQLVVPVRAALGLGHRSDLVRFRC 69
Db 80 PPGP-----RRRAGPRRRRARLARGPCGLRELEVRVSELGLGYASDETVLFRC 130
Qy 70 SGSCRRARSPhdLSLGLGALRPPGSRPVSPQCCRPTRYE-AVSFMDVNSTWRTVD 128
Db 131 AGACEAAARYVDLGLRLRRLR---RRVRAQPCCRPTAYEDEVSLDAHSRYHTVH 187
Qy 129 RLSATACGCL 138
Db 188 ELSARECACV 197

RESULT 6
US-08-742-035-7
; Sequence 7, Application US/08742035
; Patent No. 5747655
; GENERAL INFORMATION:
; APPLICANT: JOHNSON JR., EUGENE M.
; APPLICANT: MILBRANDT, JEFFREY D.
; APPLICANT: KOTZBAUER, PAUL T.
; APPLICANT: LAMPE, PATRICIA A.
; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: US
; ZIP: 63105-1817
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,035
; FILING DATE: 01-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/519,777
; FILING DATE: 28-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 953095

TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-035-7

Query Match 32.3%; Score 241; DB 1; Length 197;
Best Local Similarity 46.2%; Pred. No. 2.6e-15;
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGGRARAGGPGSRARA-AGARGCRLRSQVVPVRLGGLGHRSDLVFRFC 69
DB 80 PPGP-----RRRAGPRRRRARARLARGPCGLRELEVRVSELGLGYASDETIVLFRYC 130
QY 70 SGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYE-AVSFMDVNSTWRTVD 128
DB 131 AGACEAAARVYDLGLRLRQRRLR---RERVRAQPCCRPTAYEDEVSLDAHSRYHTVH 187
QY 129 RLSATACGCL 138
DB 188 ELSARECACV 197

RESULT 7
US-08-777-019-7
Sequence 7, Application US/08777019
Patent No. 5817622
GENERAL INFORMATION:
APPLICANT: JOHNSON JR., EUGENE M.
APPLICANT: MILBRANDT, JEFFREY D.
APPLICANT: KOTZBAUER, PAUL T.
APPLICANT: LAMPE, PATRICIA A.
TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: US
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/777,019
FILING DATE: 30-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/519,777
FILING DATE: 28-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 953095
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-777-019-7

Query Match 32.3%; Score 241; DB 2; Length 197;
Best Local Similarity 46.2%; Pred. No. 2.6e-15;
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGGRARAGGPGSRARA-AGARGCRLRSQVVPVRLGGLGHRSDLVFRFC 69
DB 80 PPGP-----RRRAGPRRRRARARLARGPCGLRELEVRVSELGLGYASDETIVLFRYC 130
QY 70 SGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYE-AVSFMDVNSTWRTVD 128
DB 131 AGACEAAARVYDLGLRLRQRRLR---RERVRAQPCCRPTAYEDEVSLDAHSRYHTVH 187
QY 129 RLSATACGCL 138
DB 188 ELSARECACV 197

RESULT 8
US-08-777-143-7
Sequence 7, Application US/08777143
Patent No. 5843914
GENERAL INFORMATION:
APPLICANT: JOHNSON JR., EUGENE M.
APPLICANT: MILBRANDT, JEFFREY D.
APPLICANT: KOTZBAUER, PAUL T.
APPLICANT: LAMPE, PATRICIA A.
TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: US
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/777,143
FILING DATE: 30-DEC-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/519,777
FILING DATE: 28-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 953095
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-777-143-7

Query Match 32.3%; Score 241; DB 2; Length 197;
Best Local Similarity 46.2%; Pred. No. 2.6e-15;
Matches 60; Conservative 10; Mismatches 46; Indels 14; Gaps 4;

QY 11 PPAPPSALPRGGRARAGGPGSRARA-AGARGCRLRSQVVPVRLGGLGHRSDLVFRFC 69
DB 80 PPGP-----RRRAGPRRRRARARLARGPCGLRELEVRVSELGLGYASDETIVLFRYC 130

TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 185:
SEQUENCE CHARACTERISTICS:
LENGTH: 156 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-931-858E-185

Query Match 31.1%; Score 232.5; DB 4; Length 156;
Best Local Similarity 43.8%; Pred. No. 1.2e-14;
Matches 53; Conservative 14; Mismatches 45; Indels 9; Gaps 2;
QY 20 RGRARAGGPGSRARAGARGLRSLVPPVRLGLGHRSDLVRFRCGSC-RRARS 78
DB 44 RGTWTPHOGNNHVRPLRALAGSCLWLSLTPVAELGLGYASEKVIIFYCAGCQPEART 103
QY 79 PHDLASLASLLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGL 138
DB 104 QHSLVLRGRG-----RAHGRPCQPTSYADVTFLDDQHHWQQLPQLSAAACGG 155
QY 139 G 139
DB 156 G 156

RESULT 14
US-08-775-414-82
Sequence 82, Application US/08775414
Patent No. 6090778
GENERAL INFORMATION:
APPLICANT: JOHNSON JR., EUGENE M.
APPLICANT: MILBRANDT, JEFFREY D.
APPLICANT: KOTZBAUER, PAUL T.
APPLICANT: LAMPE, PATRICIA A.
TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: US
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,414
FILING DATE: 31-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 965805
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-414-82

Query Match 30.9%; Score 231; DB 3; Length 142;

Best Local Similarity 46.7%; Pred. No. 1.5e-14;
Matches 57; Conservative 12; Mismatches 43; Indels 10; Gaps 4;
QY 24 AARAGGP-----GSRARAA--GARGCLRLSOLVPPVRLGLGHRSDLVRFRCGSCRRAR 77
DB 24 AARIPGPRRRAGPRRRRRRAGRGARPCGLRELVRSVSELGLGYTSDETVLFVRYCAGACEAAI 83
QY 78 SPHDLASLASLLGAGALRPPPGSRPVSPQCCRPTRYE-AVSFMDVNSTWRTVDRLSATACG 136
DB 84 RIYDLGLRLRQRRVR---RERARHPCCRPATAYEDEVSEFLDVHVSRYHTLQELSARECA 140
QY 137 CL 138
DB 141 CV 142

RESULT 15
US-08-775-414-84
Sequence 84, Application US/08775414
Patent No. 6090778
GENERAL INFORMATION:
APPLICANT: JOHNSON JR., EUGENE M.
APPLICANT: MILBRANDT, JEFFREY D.
APPLICANT: KOTZBAUER, PAUL T.
APPLICANT: LAMPE, PATRICIA A.
TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: US
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,414
FILING DATE: 31-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 965805
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 84:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-414-84

Query Match 30.9%; Score 231; DB 3; Length 150;
Best Local Similarity 46.7%; Pred. No. 1.6e-14;
Matches 57; Conservative 12; Mismatches 43; Indels 10; Gaps 4;
QY 24 AARAGGP-----GSRARAA--GARGCLRLSOLVPPVRLGLGHRSDLVRFRCGSCRRAR 77
DB 32 AARIPGPRRRAGPRRRRRRAGRGARPCGLRELVRSVSELGLGYTSDETVLFVRYCAGACEAAI 91
QY 78 SPHDLASLASLLGAGALRPPPGSRPVSPQCCRPTRYE-AVSFMDVNSTWRTVDRLSATACG 136
DB 92 RIYDLGLRLRQRRVR---RERARHPCCRPATAYEDEVSEFLDVHVSRYHTLQELSARECA 148
QY 137 CL 138

Fri Aug 17 09:56:38 2001

us-09-357-349-4.rai

Page 8

Db 149 CV 150

Search completed: August 16, 2001, 15:39:37
Job time: 67 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:44:37 ; Search time 32.99 Seconds
(without alignments)
255.433 Million cell updates/sec

Title: US-09-357-349-4
Perfect score: 139
Sequence: 1 PPQSRPAPPAPPAPPSALPR.....VNSTWRTVDRLSATACGLG 139

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 412676 seqs, 60623988 residues
Word size : 0
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 139 | 100.0 | 139 | 21 | Partial human enov |
| 2 | 139 | 100.0 | 140 | 21 | A third predicted |
| 3 | 139 | 100.0 | 140 | 21 | Amino acid sequenc |
| 4 | 139 | 100.0 | 159 | 21 | Protein-2 comprisi |
| 5 | 139 | 100.0 | 220 | 21 | Amino acid sequenc |
| 6 | 139 | 100.0 | 220 | 21 | Short splice varia |
| 7 | 139 | 100.0 | 220 | 21 | A human pre-pro-ne |
| 8 | 139 | 100.0 | 220 | 22 | Human PRO3562 prot |
| 9 | 139 | 100.0 | 228 | 21 | A human GDNF-relat |
| 10 | 139 | 100.0 | 228 | 21 | Long splice varian |
| 11 | 139 | 100.0 | 229 | 21 | Amino acid sequenc |

| | | | | | | |
|----|-----|-------|-----|----|----------|--------------------|
| 12 | 139 | 100.0 | 237 | 21 | AAV84585 | Alternatively spll |
| 13 | 139 | 100.0 | 237 | 21 | AAV92037 | Human artemin (GDN |
| 14 | 139 | 100.0 | 536 | 21 | AAV84592 | Amino acid sequenn |
| 15 | 116 | 83.5 | 116 | 21 | AAV84587 | A second predicted |
| 16 | 116 | 83.5 | 116 | 21 | AAV68712 | Amino acid sequenc |
| 17 | 113 | 81.3 | 113 | 21 | AAV84586 | A first predicted |
| 18 | 113 | 81.3 | 113 | 21 | AAV68713 | Amino acid sequenc |
| 19 | 100 | 71.9 | 113 | 21 | AAV68709 | Amino acid sequenc |
| 20 | 100 | 71.9 | 116 | 21 | AAV68708 | Amino acid sequenc |
| 21 | 100 | 71.9 | 140 | 21 | AAV68707 | Amino acid sequenc |
| 22 | 100 | 71.9 | 237 | 21 | AAV68706 | A human neublastin |
| 23 | 96 | 69.1 | 96 | 21 | AAV84589 | Amino acid sequenc |
| 24 | 65 | 46.8 | 200 | 21 | AAV68705 | Amino acid sequenc |
| 25 | 34 | 24.5 | 34 | 21 | AAV68721 | Neublastin neurotr |
| 26 | 30 | 21.6 | 160 | 21 | AAV84595 | Amino acid sequenc |
| 27 | 29 | 20.9 | 29 | 21 | AAV68722 | Neublastin neurotr |
| 28 | 28 | 20.1 | 28 | 21 | AAV68720 | Neublastin neurotr |
| 29 | 26 | 18.7 | 68 | 21 | AAV84596 | Fragment of human |
| 30 | 26 | 18.7 | 107 | 21 | AAV84598 | Protein encoded by |
| 31 | 26 | 18.7 | 123 | 21 | AAV93665 | Consensus GDNF-rel |
| 32 | 26 | 18.7 | 222 | 21 | AAV93560 | A murine GDNF-rela |
| 33 | 26 | 18.7 | 224 | 21 | AAV93558 | Amino acid sequenc |
| 34 | 26 | 18.7 | 224 | 21 | AAV84584 | A murine neublasti |
| 35 | 26 | 18.7 | 224 | 21 | AAV68714 | Neublastin neurotr |
| 36 | 17 | 12.2 | 17 | 21 | AAV68723 | Neublastin neurotr |
| 37 | 15 | 10.8 | 15 | 21 | AAV68724 | Neublastin neurotr |
| 38 | 14 | 10.1 | 14 | 21 | AAV68715 | Neublastin neurotr |
| 39 | 14 | 10.1 | 14 | 21 | AAV68716 | Neublastin neurotr |
| 40 | 14 | 10.1 | 14 | 21 | AAV68718 | Neublastin neurotr |
| 41 | 14 | 10.1 | 14 | 21 | AAV68719 | Neublastin neurotr |
| 42 | 12 | 8.6 | 12 | 21 | AAV68717 | Neublastin neurotr |
| 43 | 10 | 7.2 | 709 | 17 | AAV99801 | CR11-7 nerve prote |
| 44 | 9 | 6.5 | 72 | 14 | AAV41429 | HTLV-I gag protein |
| 45 | 9 | 6.5 | 209 | 18 | AAV26615 | Ferredoxin-HTLV-I |

ALIGNMENTS

RESULT 1
ID AAY45011 standard; Protein; 139 AA.
XX AAY45011;
XX 17-MAY-2000 (first entry)
XX Partial human enovin protein.

Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;
glial cell-line derived neurotrophic factor; GDNF; neuroprotective;
GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;
antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;
antidiarrhoeal; laxative; antileptic; neurodegenerative; Parkinson's;
Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;
endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;
dyspepsia; constipation; intestinal atony; emesis; diarrhoea;
Crohn's disease; bowel hypersensitivity.

OS Homo sapiens.

Key Peptide Location/Qualifiers
1..26
/label= Pro-sequence
/note= "Partial pro sequence of enovin"
Protein 27..139
/label= Mature Enovin
/note= "Homologous to GDNF, Neurturin and Persephin"
Modified-site 121..123
/note= "Asn is N-glycosylated"

WO200004050-A2.

PD 27-JAN-2000.
 XX
 PF 14-JUL-1999; 99WO-EP05031.
 XX
 PR 14-JUL-1998; 98GB-0015283.
 PR 12-FEB-1999; 99US-0248772.
 PR 08-JUN-1999; 99US-0327668.
 XX
 PA (JANC) JANSSEN PHARM NV.
 XX
 XX
 PI Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck LAL;
 XX
 DR WPI; 2000-182404/16.
 DR N-PSDB; AAZ50705.
 XX
 XX Novel human neurotrophic growth factor designated enovin used to treat
 PT neurological disorders, neuronal disorders, peripheral neuropathy,
 PT brain injury, nervous system disorders, prion associated and
 PT gastrointestinal diseases -
 XX
 PS Claim 1; Fig 1; 125pp; English.
 XX
 CC The present protein sequence is that of human enovin comprising
 CC partial pro sequence. Enovin (EVN) is a neurotrophic growth
 CC factor, that belongs to glial cell-line derived neurotrophic
 CC factor (GDNF) family. It binds to GDNF family receptor
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and
 CC prostate. It has nootropic, analgesic, neuroprotective, antirheumatic,
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,
 CC laxative and antiemetic activity. It can be used to treat neurological
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
 XX
 SQ Sequence 139 AA;

Query Match 100.0%; Score 139; DB 21; Length 139;
 Best Local Similarity 100.0%; Pred. No. 2.4e-115;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPQSRPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60
 Db 1 ppqsrpappppappsalprggraragpgsrraragrcrlrsqvlvpralglghrs 60
 QY 61 DELVRFRCGSGCRARRSPHDLASLLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDV 120
 Db 61 delvrfrcsgscrrarsphdlsasllgagalrpppgsrpvspccrptryeavsfmdv 120
 QY 121 NSTWRTVDRLSATACGCLG 139
 Db 121 nstwtvdrslsatagcclg 139

RESULT 2
 AAY84588
 ID AAY84588 standard; Protein: 140 AA.
 XX
 AC AAY84588;
 XX
 DT 25-JUL-2000 (first entry)
 XX
 DE A third predicted human mature artemin polypeptide.
 XX
 KW Human; artemin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;

KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.
 OS Homo sapiens.
 XX
 PN WO200018799-A1.
 XX
 PD 06-APR-2000.
 XX
 PF 29-SEP-1999; 99WO-US22604.
 XX
 PR 29-SEP-1998; 98US-0163283.
 PR 12-NOV-1998; 98US-0108148.
 PR 22-DEC-1998; 98US-0218698.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Milbrandt JD, Baloh RH;
 XX
 DR WPI; 2000-293109/25.
 DR N-PSDB; AAA12546.
 XX
 XX Isolated artemin growth factor proteins and the nucleic acids that
 PT encode them, useful for treating a range of degenerative neuronal
 PT disorders such as Parkinson's disease and Huntington's disease -
 PS
 XX Claim 4; Fig 3C; 96pp; English.

XX The present sequence represents a predicted mature human artemin growth
 CC factor protein. Artemin is a neurotrophic factor that belongs to the
 CC GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin
 CC family of growth factors and promotes differentiation, maintains mature
 CC phenotype and provides trophic support, promoting growth and survival of
 CC neurons. Artemin promotes the survival of trigeminal ganglion neurons,
 CC nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin
 CC is the only member of the GDNF family that binds to GFR-alpha (growth
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-
 CC tyrosine kinase) receptor complex and additionally, like GDNF and
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin
 CC polypeptides and polynucleotides are administered to treat peripheral
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.
 CC blastoma), multiple sclerosis, infection or enteric disease (e.g.
 CC idiopathic constipation or constipation associated with Parkinson's
 CC disease, spinal cord injury or use of opiate pain killers). They may
 CC also be used to treat a patient suffering from small cell lung carcinoma.

SQ Sequence 140 AA;

Query Match 100.0%; Score 139; DB 21; Length 140;
 Best Local Similarity 100.0%; Pred. No. 2.4e-115;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPQSRPAPPAPPALPRGGRARAGCGSRARAGARGLRSLQVPRALGLGHR 60
 Db 2 ppqsrpappppappsalprggraragpgsrraragrcrlrsqvlvpralglghrs 61
 QY 61 DELVRFRCGSGCRARRSPHDLASLLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDV 120
 Db 62 delvrfrcsgscrrarsphdlsasllgagalrpppgsrpvspccrptryeavsfmdv 121
 QY 121 NSTWRTVDRLSATACGCLG 139
 Db 122 nstwtvdrslsatagcclg 140

RESULT 3
 AAY68711
 ID AAY68711 standard; Protein: 140 AA.
 XX

PF 14-JUL-1999; 99WO-EP05031.
 XX 14-JUL-1998; 98GB-0015283.
 PR 12-FEB-1999; 99US-0248772.
 PR 08-JUN-1999; 99US-0327668.
 XX (JANC) JANSSEN PHARM NV.
 PA Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck LAL;
 PI WPI: 2000-182404/16.
 DR N-PSDB; AAZ50090.
 XX Novel human neurotrophic growth factor designated enovin used to treat
 PT neurological disorders, neuronal disorders, peripheral neuropathy,
 PT brain injury, nervous system disorders, prion associated and
 PT gastrointestinal diseases -
 XX Disclosure; Fig 3; 125pp; English.
 PS The present sequence is protein-2, comprising of partial pro
 CC sequence and mature human enovin sequence. Enovin (EVN) is a
 CC neurotrophic growth factor, that belongs to glial cell-line derived
 CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome lp31.3-32.
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and
 CC prostate. It has nootropic, analgesic, neuroprotective, antirheumatic,
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,
 CC laxative and antiemetic activity. It can be used to treat neurological
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease,
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
 XX Sequence 159 AA;
 SQ
 Query Match 100.0%; Score 139; DB 21; Length 159;
 Best Local Similarity 100.0%; Pred. No. 2.7e-115;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPOSRPAPPAPPSPALPRGGRARAGGPGSRARAGARGLRSLQVPRALGLGHR 60
 DB 21 PPGSRPAPPAPPSPALPRGGRARAGGPGSRARAGARGLRSLQVPRALGLGHR 80
 QY 61 DELVRRFCGSGCRARRSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120
 DB 81 delvrrfcsgscrrarsphdlsllgagairppgsrprvsgpcrptryeavsfmdv 140
 QY 121 NSTWRTVDRLSATACGCLG 139
 DB 141 nstwtvdrilsatagcglg 159
 RESULT 5
 ID AAY84583
 XX AAY84583 standard; Protein; 220 AA.
 AC AAY84583;
 XX 25-JUL-2000 (first entry)
 DT Amino acid sequence of a human pre-pro-arteamin polypeptide.
 DE Human; arteamin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.
 -- KW
 XX Homo sapiens.
 OS WO200018799-A1.
 PN 06-APR-2000.
 PD 29-SEP-1999; 99WO-US22604.
 XX 29-SEP-1998; 98US-0163283.
 PR 12-NOV-1998; 98US-0108148.
 PR 22-DEC-1998; 98US-0218698.
 XX (UNIW) UNIV WASHINGTON.
 PA Milbrandt JD, Baloh RH;
 PI WPI: 2000-293109/25.
 DR N-PSDB; AAA12540.
 XX Isolated arteamin growth factor proteins and the nucleic acids that
 PT encode them, useful for treating a range of degenerative neuronal
 PT disorders such as Parkinson's disease and Huntington's disease -
 XX Claim 5; Fig 1B; 96pp; English.
 PS The present sequence represents a pre-pro- arteamin growth factor protein.
 CC Arteamin is a neurotrophic factor that belongs to the GDNF (glial cell
 CC line-derived neurotrophic factor)/neurturin/persephin family of growth
 CC factors and promotes differentiation, maintains mature phenotype and
 CC provides trophic support, promoting growth and survival of neurons.
 CC Arteamin promotes the survival of trigeminal ganglion neurons, nodose
 CC ganglion neurons, superior cervical ganglion neurons and tyrosine-
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Arteamin
 CC is the only member of the GDNF family that binds to GFR-alpha (growth
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-
 CC tyrosine kinase) receptor complex and additionally, like GDNF and
 CC neurturin, arteamin also binds to and activates GFRalpha/RET. Arteamin
 CC polypeptides and polynucleotides are administered to treat peripheral
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.
 CC idiopathic constipation or constipation associated with Parkinson's
 CC disease, spinal cord injury or use of opiate pain killers). They may
 CC also be used to treat a patient suffering from small cell lung carcinoma.
 XX Sequence 220 AA;
 SQ
 Query Match 100.0%; Score 139; DB 21; Length 220;
 Best Local Similarity 100.0%; Pred. No. 3.6e-115;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPOSRPAPPAPPSPALPRGGRARAGGPGSRARAGARGLRSLQVPRALGLGHR 60
 DB 82 PPGSRPAPPAPPSPALPRGGRARAGGPGSRARAGARGLRSLQVPRALGLGHR 141
 QY 61 DELVRRFCGSGCRARRSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120
 DB 142 delvrrfcsgscrrarsphdlsllgagairppgsrprvsgpcrptryeavsfmdv 201
 QY 121 NSTWRTVDRLSATACGCLG 139
 DB 202 nstwtvdrilsatagcglg 220
 RESULT 6
 ID AAY44776
 XX AAY44776 standard; Protein; 220 AA.
 AC AAY44776;
 XX

06-JUL-1998; 98DK-0000904.
 09-JUL-1998; 98US-0092229.
 19-AUG-1998; 98DK-0001048.
 25-AUG-1998; 98US-0097774.
 06-OCT-1998; 98DK-0001265.
 13-OCT-1998; 98US-0103908.
 02-JUL-1999; 99US-0347613.
 (NEUR-) NEUROSEARCH AS.
 Johansen TE, Blom N, Hansen C;
 WPI: 2000-171013/15.
 N-PSDB; AA260563.
 New isolated polypeptides, used for treating e.g. neurodegenerative disease or disorder, neuronal damage or neuronal disorder of the peripheral nervous system, the medulla or the spinal cord -
 Claim 14; Page 97; 106pp; English.
 The present sequence represents a neurotrophic factor designated neublastin. Neublastin is a member of the glial cell line-derived neurotrophic factor sub-class of the transforming growth factor-beta superfamily of neurotrophic factors. Neublastin exhibits high affinity for the GFR-alpha3-RRT receptor complex. The polypeptides can be used for treating a neurodegenerative disease or disorder, cerebral ischemic neuronal damage, traumatic brain injury, peripheral neuropathy, Alzheimer's disease, Huntington's disease, Parkinson's disease, Parkinson-Plus syndromes, progressive Supranuclear Palsy, Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian parkinsonism dementia complex, amyotrophic lateral sclerosis, memory impairment, or a neuronal disorder of the peripheral nervous system, the medulla or the spinal cord. They can also be used for treating various neuropathies. They can also be used for treating ischemic stroke, acute brain injury, acute spinal cord injury, nervous system tumours, multiple sclerosis, exposure to neurotoxins, metabolic diseases such as diabetes or renal dysfunctions and damage caused by infectious agents, or various disorders in the eye.

Sequence 220 AA;

Query Match 100.0%; Score 139; DB 21; Length 220;
 Best Local Similarity 100.0%; Pred. No. 3.6e-115;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPOPSRPAPPAPPALPRGGRAARAGGPGSRARAGCRLRSQVLPVRAIGLGHRS 60
 Db 82 PPOPSRPAPPAPPALPRGGRAARAGGPGSRARAGCRLRSQVLPVRAIGLGHRS 141
 QY 61 DELVRFRCGSCRRARSPHDLASLLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDV 120
 Db 142 delvrfrcgscrrarsphdlsllgagallrpppgsrpvsqpcrptryeavsfmdv 201
 QY 121 NSTWRTVDRLSATACGCLG 139
 Db 202 nstwtvdrslsatcgclg 220

RESULT 8
 ID AAB50978 standard; Protein; 220 AA.
 AC AAB50978;
 DT 21-MAR-2001 (first entry)
 XX Human PRO3562 protein.
 DE Human; PRO; cytostatic; neuroprotective; respiratory general;
 KW antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant;
 -- KW PRO agonist; cancer; inflammatory disorder; immunological disorder.

XX Homo sapiens.
 OS WO200073348-A2.
 PN 07-DEC-2000.
 XX 30-MAY-2000; 2000WO-US14941.
 XX 02-JUN-1999; 99WO-US12252.
 PR 22-JUN-1999; 99US-0140650.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 29-OCT-1999; 99US-0162506.
 PR 30-NOV-1999; 99WO-US28313.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 16-DEC-1999; 99WO-US30095.
 PR 06-DEC-1999; 99WO-US30999.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 03-MAR-2000; 2000WO-US05841.
 PR 10-MAR-2000; 2000US-0187202.
 PR 15-MAR-2000; 2000WO-US06319.
 PR 30-MAR-2000; 2000WO-US06884.
 PR 17-MAY-2000; 2000WO-US13705.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC;
 PI Shelton DL, Smith V, Watanabe CK, Wood WI;
 XX WPI: 2001-016509/02.
 DR N-PSDB; AAC91580.

XX Twenty eight nucleic acids encoding PRO polypeptides which are useful
 PT for treating various tumors, e.g. breast cancer, and other
 PT inflammatory, angiogenic and immunological disorders -
 XX Claim 31; Fig 56; 188pp; English.

XX The present sequence is one of twenty eight novel PRO polypeptides. The
 CC PRO polypeptides and their agonists, including antibodies, peptides, and
 CC small molecule agonists, may be used to treat various tumors, e.g.,
 CC cancers such as breast cancer, ovarian cancer, renal cancer, colorectal
 CC cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,
 CC central nervous system cancer, melanoma or leukaemia. They are also
 CC useful for treating other disorders such as neuronal, glial, astrocytal,
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal and
 CC blastocoealic disorders, and inflammatory, angiogenic and immunological
 CC disorders.

XX Sequence 220 AA;

Query Match 100.0%; Score 139; DB 22; Length 220;
 Best Local Similarity 100.0%; Pred. No. 3.6e-115;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPOPSRPAPPAPPALPRGGRAARAGGPGSRARAGCRLRSQVLPVRAIGLGHRS 60
 Db 82 PPOPSRPAPPAPPAPPALPRGGRAARAGGPGSRARAGCRLRSQVLPVRAIGLGHRS 141
 QY 61 DELVRFRCGSCRRARSPHDLASLLGAGALRPPPGSRPVSPQCCRPTRYEAVSFMDV 120
 Db 142 delvrfrcgscrrarsphdlsllgagallrpppgsrpvsqpcrptryeavsfmdv 201
 QY 121 NSTWRTVDRLSATACGCLG 139

Db 202 nstwtvdrisatcacgclg 220
|||||

RESULT 9

AA93559
ID AAY93559 standard; Protein; 228 AA.

XX AC AAY93559;

XX DT 25-SEP-2000 (first entry)

XX DE A human GDNF-related neurotrophic factor 4 (GRNF4).

XX KW GDNF; glial cell line-derived neurotrophic factor; GFRalpha-3;
KW GDNF-related neurotrophic factor 4; GRNF4; GDNF family receptor-alpha-3;
KW Parkinson's disease; Alzheimer's disease; amyotrophic lateral sclerosis;
KW incontinence; bone loss; osteoporosis; osteogenesis imperfecta;
KW hypercalcemia; nerve damage; stroke; cancer; dideoxycytidine; AIDS;
KW chronic metabolic disease; renal dysfunction.

XX OS Homo sapiens.

XX PN W0200034475-A2.

XX PD 15-JUN-2000.

XX PF 08-DEC-1999; 99WO-US28975.

XX PR 09-DEC-1998; 98US-0111626.

XX PA (AMGE-) AMGEN INC.

XX PI Simonet WS, Asuncion EJ, Min H, Jing S;

XX DR WPI; 2000-423421/36.

XX DR N-PSDB; AAA46615.

XX PT New glial cell line-derived neurotrophic factor-related neurotrophic
PT factor 4 useful for treating neurodegenerative disease such as
PT Parkinson's disease and for treating nerve damage caused by physical
PT injury and other metabolic diseases

XX PS Claim 1; Fig 7; 135pp; English.

XX CC The present sequence represents a human GDNF (glial cell line-derived
CC neurotrophic factor)-related neurotrophic factor 4 (GRNF4) protein.
CC The GRNF4 polypeptide is capable of binding a GDNF family
CC receptor-alpha-3 (GFRalpha-3). The GRNF4 polynucleotides may be
CC used for in vitro GRNF4 protein production as well as in cell therapy
CC or gene therapy applications. GRNF4 protein product may be used in
CC treating, Parkinson's disease, Alzheimer's disease, amyotrophic
CC lateral sclerosis, incontinence, diseases associated with bone loss
CC (e.g. osteoporosis, osteogenesis imperfecta or hypercalcemia of
CC malignancy). GRNF4 protein products may also be used in the treatment
CC of nerve damage which may occur to one or more types of nerve cells by
CC physical injury, which causes the degeneration of the axonal processes
CC and/or nerve cell bodies near the site of injury, temporary or permanent
CC cessation of blood flow to parts of the nervous system, as in stroke,
CC intentional or accidental exposure to neurotoxins, for e.g.
CC chemotherapeutic agents for the treatment of cancer or dideoxycytidine
CC for the treatment of AIDS, chronic metabolic diseases, including
CC diabetes or renal dysfunction.

XX SQ Sequence 228 AA;

Query Match 100.0%; Score 139; DB 21; Length 228;
Best Local Similarity 100.0%; Pred. No. 3.7e-115;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPQSPRAPPAPPSPALPGRRAGGPGSRARAGCRLRSQLPVVRALGLGHS 60
|||||

Db 90 PPQSPRAPPAPPSPALPGRRAGGPGSRARAGCRLRSQLPVVRALGLGHS 149
QY 61 DELVRFECSCRRARPHDLSLASLGLAGALRPPGSRPVSPCCRPTRVEAVSFMDV 120
|||||
Db 150 delvrfecscrrarsphdlsiaslligagalprrpsrpsqccrptryeavsfmdv 209
QY 121 NSTWRTVDRLSATACGCLG 139
|||||
Db 210 nstwtvdrisatcacgclg 228

RESULT 10

AA944775
ID AAY44775 standard; Protein; 228 AA.

XX AC AAY44775;

XX DT 17-MAY-2000 (first entry)

XX DE Long splice variant of human Enovin.

XX KW Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;

XX KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;

XX KW GDNF family receptor alpha-3; GFR alpha 3; nocotropic; analgesic;

XX KW antirheumatic; cerebroprotective; antiparkinsonian; antinflammatory;

XX KW antidarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;

XX KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;

XX KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;

XX KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;

XX KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..47

FT Peptide /label= Signal_Peptide

FT Peptide 48..115

FT Peptide /label= Pro_sequence

FT Misc-difference 89..228

FT Protein /note= "This region has been claimed specifically"

FT Protein 116..228

FT Modified-site /label= Mature_Enovin

FT Modified-site /note= "Homologous to GDNF, Neurturin and Persephin"

FT Modified-site 210..212

FT Modified-site /note= "Asn is N-glycosylated"

FT Misc-difference 131

FT Misc-difference /note= "Conserved residue characteristic of Transforming

FT Misc-difference 158

FT Misc-difference /note= "Conserved residue characteristic of Transforming

FT Misc-difference 162

FT Misc-difference /note= "Conserved residue characteristic of Transforming

FT Misc-difference 195

FT Misc-difference /note= "Conserved residue characteristic of Transforming

FT Misc-difference 196

FT Misc-difference /note= "Conserved residue characteristic of Transforming

FT Misc-difference 224

FT Misc-difference /note= "Conserved residue characteristic of Transforming

FT Misc-difference 226

FT Misc-difference /note= "Conserved residue characteristic of Transforming

FT Misc-difference 228

FT Misc-difference /note= "Conserved residue characteristic of Transforming

FT Misc-difference 228

FT Misc-difference /note= "Conserved residue characteristic of Transforming

PN W0200004050-A2.

XX PD 27-JAN-2000.

XX PF 14-JUL-1999; 99WO-EP05031.

PR 14-JUL-1998; 98GB-0015283.
 PR 12-FEB-1999; 99US-0248772.
 PR 08-JUN-1999; 99US-0327668.
 XX (JANC) JANSSEN PHARM NV.
 XX Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck LAL;
 PI WPI: 2000-182404/16.
 DR N-PSDB; AAZ50091.
 XX Novel human neurotrophic growth factor designated enovin used to treat
 PT neurological disorders, neuronal disorders, peripheral neuropathy,
 PT brain injury, nervous system disorders, prion associated and
 PT gastrointestinal diseases
 XX Claim 11; Fig 23; 125pp; English.
 XX The present sequence is a long splice variant of human Enovin (EVN). EVN
 CC is a neurotrophic growth factor, that belongs to glial cell-line derived
 CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and
 CC prostate. It has nootropic, analgesic, neuroprotective, antirheumatic,
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,
 CC laxative and antileptic activity. It can be used to treat neurological
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease.
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
 CC EVN polynucleotide can be used in gene therapy.
 XX Sequence 228 AA:
 SQ
 Query Match 100.0%; Score 139; DB 21; Length 228;
 Best Local Similarity 100.0%; Pred. No. 3.7e-115;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPOPSRPPAPPAPPALPRGGAARAGGPGSRARAGCRLRSQVPRALGLGHS 60
 DB 90 PPOPSRPPAPPAPPALPRGGAARAGGPGSRARAGCRLRSQVPRALGLGHS 149
 QY 61 DELVFRFCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120
 DB 150 DELVFRFCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 209
 QY 121 NSTWRTVDRLSATACGCLG 139
 DB 210 NSTWRTVDRLSATACGCLG 228
 RESULT 11
 ID AAY84582 standard; Protein; 229 AA.
 XX AAY84582;
 XX 25-JUL-2000 (first entry)
 XX Amino acid sequence of a human artemin polypeptide.
 DE Human; artemin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.
 XX Homo sapiens.
 OS

XX WO200018799-A1.
 XX 06-APR-2000.
 XX 29-SEP-1999; 99WO-US22604.
 XX 29-SEP-1998; 98US-0163283.
 PR 12-NOV-1998; 98US-0108148.
 PR 22-DEC-1998; 98US-0218698.
 XX (UNIW) UNIV WASHINGTON.
 XX Milbrandt JD, Baloh RH;
 PI WPI: 2000-293109/25.
 DR N-PSDB; AAA12539.
 XX Isolated artemin growth factor proteins and the nucleic acids that
 PT encode them, useful for treating a range of degenerative neuronal
 PT disorders such as Parkinson's disease and Huntington's disease -
 XX Disclosure; Fig 1A; 96pp; English.
 XX The present sequence represents a human artemin growth factor protein.
 CC Artemin is a neurotrophic factor that belongs to the GDNF (glial cell
 CC line-derived neurotrophic factor)/neurturin/persephin family of growth
 CC factors and promotes differentiation, maintains mature phenotype and
 CC provides trophic support, promoting growth and survival of neurons.
 CC Artemin promotes the survival of trigeminal ganglion neurons, nodose
 CC ganglion neurons, superior cervical ganglion neurons and tyrosine-
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin
 CC is the only member of the GDNF family that binds to GFR-alpha (growth
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-
 CC tyrosine kinase) receptor complex and additionally, like GDNF and
 CC neurturin, artemin also binds to and activates GFRalpha1/RET. Artemin
 CC polypeptides and polynucleotides are administered to treat peripheral
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.
 CC blastoma), multiple sclerosis, infection or enteric disease (e.g.
 CC idiopathic constipation or constipation associated with Parkinson's
 CC disease, spinal cord injury or use of opiate pain killers). They may
 CC also be used to treat a patient suffering from small cell lung carcinoma.
 XX Sequence 229 AA:
 SQ
 Query Match 100.0%; Score 139; DB 21; Length 229;
 Best Local Similarity 100.0%; Pred. No. 3.7e-115;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPOPSRPPAPPAPPALPRGGAARAGGPGSRARAGCRLRSQVPRALGLGHS 60
 DB 77 PPOPSRPPAPPAPPALPRGGAARAGGPGSRARAGCRLRSQVPRALGLGHS 136
 QY 61 DELVFRFCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 120
 DB 137 DELVFRFCGSCRRARSPHDLASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDV 196
 QY 121 NSTWRTVDRLSATACGCLG 139
 DB 197 NSTWRTVDRLSATACGCLG 215
 RESULT 12
 ID AAY84585 standard; Protein; 237 AA.
 XX AAY84585;
 XX 25-JUL-2000 (first entry)
 XX Amino acid sequence of a human artemin polypeptide.
 DE Human; artemin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.
 XX Homo sapiens.
 OS

Alternatively spliced human artemin polypeptide.

Human; artemin; growth factor; neurotrophic factor; trophic support; neuron; trigeminal ganglion neuron; nodose ganglion neuron; superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease; peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke; Parkinson's disease; Huntington's disease; acute brain injury; acute spinal cord injury; nervous system tumor; blastoma; multiple sclerosis; infection; enteric disease; idiopathic constipation; Parkinson's disease; small cell lung carcinoma.

Homo sapiens.

WO200018799-A1.

06-APR-2000.

29-SEP-1999; 99WO-US22604.

29-SEP-1998; 98US-0163283.

12-NOV-1998; 98US-0108148.

22-DEC-1998; 98US-0218698.

(UNIW) UNIV WASHINGTON.

Milbrandt JD, Baloh RH;

WPI; 2000-293109/25.

N-PSDB; AAA12542.

Isolated artemin growth factor proteins and the nucleic acids that encode them, useful for treating a range of degenerative neuronal disorders such as Parkinson's disease and Huntington's disease.

Claim 5; Fig 1D; 96pp; English.

The present sequence is encoded by an alternatively spliced human artemin growth factor cDNA. Artemin is a neurotrophic factor that belongs to the GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin family of growth factors and promotes differentiation, maintains mature phenotype and provides trophic support, promoting growth and survival of neurons. Artemin promotes the survival of trigeminal ganglion neurons, nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin is the only member of the GDNF family that binds to GFR-alpha (growth factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-tyrosine kinase) receptor complex and additionally, like GDNF and neurturin, artemin also binds to and activates GFRalpha1/RET. Artemin polypeptides and polynucleotides are administered to treat peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, acute brain injury, acute spinal cord injury, a nervous system tumor (e.g. blastoma), multiple sclerosis, infection or enteric disease (e.g. idiopathic constipation or constipation associated with Parkinson's disease, spinal cord injury or use of opiate pain killers). They may also be used to treat a patient suffering from small cell lung carcinoma.

Sequence 237 AA;

Query Match 100.0%; Score 139; DB 21; Length 237;
Best Local Similarity 100.0%; Pred. No. 3.8e-115;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 PQQSPAPPPAPPSPALPRGRRARAGGPGSRAAGARCRRLSOLVPPVRLGLHRS 60
99 PQQSPAPPPAPPSPALPRGRRARAGGPGSRAAGARCRRLSOLVPPVRLGLHRS 158
61 DELVRFECGSCRRARSPhDLASLLGAGALRPPGSRPVSQPCRPTRYEAVSFMDV 120
159 delvrfecgscrrarsphdlasllgagallrppgsrpsvqpcrptryeavsfmdv 218

QY 121 NSTWRTVDRLSATACGCLG 139
Db 219 nstwtvdrslsatacgclg 237
|||||
RESULT 13
AAY92037
ID AAY92037 standard; Protein; 237 AA.
XX
AC AAY92037;
XX
DT 19-JUL-2000 (first entry)
XX
DE Human artemin (GDNF) subunit.
XX
KW human artemin; GDNF; CKGF; mutant; cystine knot growth factor;
KW hairpin loop; infertility.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..143 /note= "optionally mutated to increase electrostatic
FT interaction between beta hairpin structure and
FT a receptor"
FT Domain 144..163
FT /label= beta_hairpin_loop_1
FT /note= "mutant optionally comprises one or more
FT substitutions in these residues"
FT Misc-difference 164..208
FT /note= "optionally mutated to increase electrostatic
FT interaction between beta hairpin structure and
FT a receptor"
FT Domain 209..229
FT /label= beta_hairpin_loop_3
FT /note= "mutant optionally comprises one or more
FT substitutions in these residues"
FT Misc-difference 230..237
FT /note= "optionally mutated to increase electrostatic
FT interaction between beta hairpin structure and
FT a receptor"
XX
PN WO200017360-A1.
XX
PD 30-MAR-2000.
XX
PF 19-MAR-1999; 99WO-US05908.
XX
PR 22-SEP-1998; 98WO-US19772.
XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Weintraub BD, Szkuudlinski MW;
XX
DR WPI; 2000-283585/24.
XX
PT New mutant cystine knot growth factor proteins comprising one or more
PT mutant subunits, useful for treating or preventing diseases e.g.
PT hypothyroidism and thyroid cancer
XX
PS Claim 594; Page 314; 320pp; English.
XX
CC This is the wild type human artemin (GDNF).
CC Mutants comprise at least one electrostatic charge altering mutation in a
CC beta hairpin loop, resulting in increased bioactivity.
CC Mutant cystine knot growth factor (CKGF) proteins comprising one or more
CC mutant subunits and having novel properties or improved pharmacological
CC properties, compared to wild type CKGFs, are claimed. The CKGF
CC superfamily comprises at least four families of growth factors: the
CC glycoprotein hormones, the platelet-derived growth factor (PDGF) family,
CC the neurotrophins and the transforming growth factor-beta family; the
CC families are known to be structurally similar (especially comprising the
CC cystine knot topology) and it was shown that mutations at certain

XX (UNIW) UNIV WASHINGTON.
 XX PI Milbrandt JD, Baloh RH;
 XX DR WPI; 2000-293109/25.
 XX DR N-PSDB; AAA12545.
 XX
 PT Isolated artemin growth factor proteins and the nucleic acids that
 PT encode them, useful for treating a range of degenerative neuronal
 PT disorders such as Parkinson's disease and Huntington's disease -
 XX
 PS Claim 4; Fig 3B; 96pp; English.
 XX
 CC The present sequence represents a predicted mature human artemin growth
 CC factor protein. Artemin is a neurotrophic factor that belongs to the
 CC GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin
 CC family of growth factors and promotes differentiation, maintains mature
 CC phenotype and provides trophic support, promoting growth and survival of
 CC neurons. Artemin promotes the survival of trigeminal ganglion neurons,
 CC nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin
 CC is the only member of the GDNF family that binds to GFR-alpha (growth
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-
 CC tyrosine kinase) receptor complex and additionally, like GDNF and
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin
 CC polypeptides and polynucleotides are administered to treat peripheral
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.
 CC idiopathic constipation or constipation associated with Parkinson's
 CC disease, spinal cord injury or use of opiate pain killers). They may
 CC also be used to treat a patient suffering from small cell lung carcinoma.
 XX
 SQ Sequence 116 AA;

Query Match 83.5%; Score 116; DB 21; Length 116;
 Best Local Similarity 100.0%; Pred. No. 4.2e-95;
 Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 24 AARAGGGSRRARAGCRLRSQVLPVRLGLGHRSDLVRFRCGSCRRARSPHDL 83
 DB 1 AARAGGGSRRARAGCRLRSQVLPVRLGLGHRSDLVRFRCGSCRRARSPHDL 60
 QY 84 LASLLGAGALRPPGSRPVSPCCRPTRYEAVSFMDVNSTWTVDFLSATACGCLG 139
 DB 61 lasllgagalrppgsrpvspccrptryeavsfmdvnsntwrtvdrfsatagcglg 116

Search completed: August 16, 2001, 15:44:38
 Job time: 298 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:38:30 ; Search time 33.13 Seconds
(without alignments)
206.777 Million cell updates/sec

Title: US-09-357-349-3

Perfect score: 601

Sequence: 1 AGPGSRARAAGRGRLRS.....VNSTWRTVDRLSATAGCCLG 113

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 601 | 100.0 | 113 | 21 | AAV84586 |
| 2 | 601 | 100.0 | 113 | 21 | AAV68713 |
| 3 | 601 | 100.0 | 116 | 21 | AAV84587 |
| 4 | 601 | 100.0 | 116 | 21 | AAV68712 |
| 5 | 601 | 100.0 | 139 | 21 | AAV45011 |
| 6 | 601 | 100.0 | 140 | 21 | AAV84588 |
| 7 | 601 | 100.0 | 140 | 21 | AAV68711 |
| 8 | 601 | 100.0 | 159 | 21 | AAV44774 |
| 9 | 601 | 100.0 | 220 | 21 | AAV84583 |
| 10 | 601 | 100.0 | 220 | 21 | AAV44776 |
| 11 | 601 | 100.0 | 220 | 21 | AAV68710 |

| | | | | | | |
|----|-------|-------|-----|----|----------|--------------------|
| 12 | 601 | 100.0 | 220 | 22 | AAV50978 | Human PRO3562 prot |
| 13 | 601 | 100.0 | 228 | 21 | AAV93559 | A human GDNF-relat |
| 14 | 601 | 100.0 | 228 | 21 | AAV44775 | Long splice varian |
| 15 | 601 | 100.0 | 229 | 21 | AAV84582 | Amino acid sequenc |
| 16 | 601 | 100.0 | 237 | 21 | AAV84585 | Alternatively spli |
| 17 | 601 | 100.0 | 237 | 21 | AAV92037 | Human artemin (GDN |
| 18 | 601 | 100.0 | 536 | 21 | AAV84592 | Amino acid sequenc |
| 19 | 589 | 98.0 | 113 | 21 | AAV68709 | Amino acid sequenc |
| 20 | 589 | 98.0 | 116 | 21 | AAV68708 | Amino acid sequenc |
| 21 | 589 | 98.0 | 140 | 21 | AAV68707 | Amino acid sequenc |
| 22 | 588 | 97.8 | 237 | 21 | AAV68706 | A human neublastin |
| 23 | 536.5 | 89.3 | 200 | 21 | AAV68705 | Amino acid sequenc |
| 24 | 528 | 87.9 | 123 | 21 | AAV93665 | Protein encoded by |
| 25 | 528 | 87.9 | 224 | 21 | AAV93558 | A murine GDNF-rela |
| 26 | 528 | 87.9 | 224 | 21 | AAV84584 | Amino acid sequenc |
| 27 | 528 | 87.9 | 224 | 21 | AAV68714 | A murine neublasti |
| 28 | 517 | 86.0 | 222 | 21 | AAV93560 | Consensus GDNF-rel |
| 29 | 515 | 85.7 | 96 | 21 | AAV84589 | Amino acid sequenc |
| 30 | 458 | 76.2 | 160 | 21 | AAV84595 | Amino acid sequenc |
| 31 | 237 | 39.4 | 156 | 20 | AAV16727 | Human pre-pro pers |
| 32 | 237 | 39.4 | 156 | 21 | AAV92038 | Human glial cell d |
| 33 | 232 | 38.6 | 197 | 18 | AAV13716 | Human pre-pro-neur |
| 34 | 232 | 38.6 | 197 | 20 | AAV16637 | WO914235 Seq ID N |
| 35 | 232 | 38.6 | 197 | 20 | AAV00863 | Human pre-pro-neur |
| 36 | 232 | 38.6 | 197 | 21 | AAV92014 | Human neurturin. |
| 37 | 228.5 | 38.0 | 134 | 18 | AAV30067 | Mouse persephin. |
| 38 | 228.5 | 38.0 | 134 | 20 | AAV16663 | WO914235 Seq ID N |
| 39 | 228.5 | 38.0 | 142 | 20 | AAV16681 | WO914235 Seq ID N |
| 40 | 228.5 | 38.0 | 156 | 20 | AAV16721 | Murine pre-pro per |
| 41 | 228.5 | 38.0 | 185 | 18 | AAV26680 | Mature mouse perse |
| 42 | 228.5 | 38.0 | 185 | 20 | AAV16652 | WO914235 Seq ID N |
| 43 | 228 | 37.9 | 96 | 18 | AAV30075 | Persephin-neurturi |
| 44 | 228 | 37.9 | 96 | 20 | AAV16695 | WO914235 Seq ID N |
| 45 | 227.5 | 37.9 | 96 | 18 | AAV30066 | Mature mouse perse |

ALIGNMENTS

RESULT 1
AAV84586
ID AAV84586 standard; Protein; 113 AA.
XX AC AAV84586;
XX DT 25-JUL-2000 (first entry)
XX DE A first predicted human mature artemin polypeptide.

XX Human; artemin; growth factor; neurotrophic factor; trophic support;
KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
KW Parkinson's disease; Huntington's disease; acute brain injury;
KW acute spinal cord injury; nervous system tumour; blastoma;
KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
KW Parkinson's disease; small cell lung carcinoma.

OS Homo sapiens.
PN WO200018799-A1.
XX 06-APR-2000.

XX 29-SEP-1999; 99WO-US22604.

XX 29-SEP-1998; 98US-0163283.

XX 12-NOV-1998; 98US-0108148.

XX 22-DEC-1998; 98US-0218698.

XX (UNIW) UNIV WASHINGTON.

XX Milbrandt JD, Baloh RH;

XX WPI; 2000-293109/25.
 DR N-PSDB; AAA12543.
 XX
 PT Isolated artemin growth factor proteins and the nucleic acids that
 PT encode them, useful for treating a range of degenerative neuronal
 PT disorders such as Parkinson's disease and Huntington's disease -
 XX
 XX
 PS Claim 4; Fig 3A; 96pp; English.
 XX
 XX The present sequence represents a predicted mature human artemin growth
 CC factor protein. Artemin is a neurotrophic factor that belongs to the
 CC GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin
 CC family of growth factors and promotes differentiation, maintains mature
 CC phenotype and provides trophic support, promoting growth and survival of
 CC neurons. Artemin promotes the survival of trigeminal ganglion neurons,
 CC nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin
 CC is the only member of the GDNF family that binds to GFR-alpha (growth
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-
 CC tyrosine kinase) receptor complex and additionally, like GDNF and
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin
 CC polypeptides and polynucleotides are administered to treat peripheral
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.
 CC idiopathic constipation or constipation associated with Parkinson's
 CC disease, spinal cord injury or use of opiate pain killers). They may
 CC also be used to treat a patient suffering from small cell lung carcinoma.
 XX
 XX Sequence 113 AA;

Query Match 100.0%; Score 601; DB 21; Length 113;
 Best Local Similarity 100.0%; Pred. No. 1.2e-56;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGPGSRARAAGARGCRLRSQVPRALGLHRSDELVRFCSGCRARRSPHDLISLAS 60
 DB 1 agpgsraraagargcrlrsqvpvralglhrrdelvrfcsgcrarrsphdlislas 60
 QY 61 LLGAGALRPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 DB 61 llgagalrppgsrpsvpsccrptryeavsfmdvnsdwtvdrilsatacgcg 113

RESULT 2
 AAY68713
 ID AAY68713 standard; Protein; 113 AA.
 XX
 AC AAY68713;
 XX
 DT 05-MAY-2000 (first entry)
 XX
 XX Amino acid sequence of a neublastin neurotrophic factor variant NBN113.
 DE
 XX Neurotrophic factor; neublastin; neurodegenerative disease;
 KW cerebral ischemic neuronal damage; traumatic brain injury;
 KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;
 KW Parkinson's disease; Parkinson-Plus syndrome;
 KW progressive Supranuclear Palsy; Olivopontocerebellar atrophy;
 KW Shy-Drager Syndrome; Guamanian parkinsonism dementia complex;
 KW amyotrophic lateral sclerosis; memory impairment; neuronal disorder;
 KW neuropathy; ischemic stroke; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; multiple sclerosis;
 KW neurotoxin exposure; metabolic disease; diabetes; renal dysfunction;
 KW eye disorder.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 95
 FT

/note= "glycosylated residue"
 WO200001815-A2.
 13-JAN-2000.
 05-JUL-1999; 99WO-DK00384.
 06-JUL-1998; 98DK-0000904.
 09-JUL-1998; 98US-0002229.
 19-AUG-1998; 98DK-0001048.
 25-AUG-1998; 98US-0007774.
 06-OCT-1998; 98DK-0001265.
 13-OCT-1998; 98US-0103908.
 02-JUL-1999; 99US-0347613.
 (NEUR-) NEUROSEARCH AS.
 Johansen TE, Blom N, Hansen C;
 WPI; 2000-171013/15.
 New isolated polypeptides, used for treating e.g. neurodegenerative
 disease or disorder, neuronal damage or neuronal disorder of the
 peripheral nervous system, the medulla or the spinal cord -
 Claim 14; Page 99; 106pp; English.

The present sequence represents a variant of a neurotrophic factor
 designated neublastin. Neublastin is a member of the glial cell line-
 derived neurotrophic factor sub-class of the transforming growth factor-
 beta superfamily of neurotrophic factors. Neublastin exhibits high
 affinity for the GFR-alpha3-RET receptor complex. The polypeptides can
 be used for treating a neurodegenerative disease or disorder, cerebral
 ischemic neuronal damage, traumatic brain injury, peripheral neuropathy,
 Alzheimer's disease, Huntington's disease, Parkinson's disease,
 Parkinson-Plus syndrome, progressive Supranuclear Palsy,
 Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian
 parkinsonism dementia complex, amyotrophic lateral sclerosis, memory
 impairment, or a neuronal disorder of the peripheral nervous system,
 the medulla or the spinal cord. They can also be used for treating
 various neuropathies. They can also be used for treating ischemic stroke,
 acute brain injury, acute spinal cord injury, nervous system tumours,
 multiple sclerosis, exposure to neurotoxins, metabolic diseases such as
 diabetes or renal dysfunctions and damage caused by infectious agents,
 or various disorders in the eye.

Query Match 100.0%; Score 601; DB 21; Length 113;
 Best Local Similarity 100.0%; Pred. No. 1.2e-56;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGPGSRARAAGARGCRLRSQVPRALGLHRSDELVRFCSGCRARRSPHDLISLAS 60
 DB 1 agpgsraraagargcrlrsqvpvralglhrrdelvrfcsgcrarrsphdlislas 60
 QY 61 LLGAGALRPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 DB 61 llgagalrppgsrpsvpsccrptryeavsfmdvnsdwtvdrilsatacgcg 113

RESULT 3
 AAY84587
 ID AAY84587 standard; Protein; 116 AA.
 XX
 AC AAY84587;
 XX
 DT 25-JUL-2000 (first entry)
 XX
 XX A second predicted human mature artemin polypeptide.

factor protein. Artemin is a neurotrophic factor that belongs to the GDNF (glial cell line-derived neurotrophic factor)/neurturin/perserpin family of growth factors and promotes differentiation, maintains mature phenotype and provides trophic support, promoting growth and survival of neurons. Artemin promotes the survival of trigeminal ganglion neurons, nodose ganglion neurons, superior cervical ganglion neurons and tyrosine hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin is the only member of the GDNF family that binds to GFR-alpha (growth factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-tyrosine kinase) receptor complex and additionally, like GDNF and neurturin, artemin also binds to and activates GFRalpha/RET. Artemin polypeptides and polynucleotides are administered to treat peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, acute brain injury, acute spinal cord injury, a nervous system tumour (e.g. blastomas), multiple sclerosis, infection or enteric disease (e.g. idiopathic constipation or constipation associated with Parkinson's disease, spinal cord injury or use of opiate pain killers). They may also be used to treat a patient suffering from small cell lung carcinoma.

Sequence 140 AA;

Query Match 100.0%; Score 601; DB 21; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.5e-56;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGPGSARAAGARGCLRLSOLVPRALGHRSDRLVRFRCSCGRRARSPHDLSLAS 60
DB 28 agpgsaraagargclrlsqvpralglhrsdrlvrfrcsgrrarsphdlslas 87
QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
DB 88 llgagalrpppgsrpvspccrptryeavsfmdvnsrtvdrisatagclg 140

RESULT 7

AY68711
ID AAY68711 standard; Protein; 140 AA.

AC AAY68711;

DT 05-MAY-2000 (first entry)

DE Amino acid sequence of a neublastin neurotrophic factor variant NBN140.

KW Neurotrophic factor; neublastin; neurodegenerative disease;
KW cerebral ischemic neuronal damage; traumatic brain injury;
KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;
KW Parkinson's disease; Parkinson-Plus syndrome;
KW progressive Supranuclear Palsy; Olivopontocerebellar atrophy;
KW Shy-Drager Syndrome; Guamanian parkinsonism dementia complex;
KW amyotrophic lateral sclerosis; memory impairment; neuronal disorder;
KW neuropathy; ischemic stroke; acute brain injury;
KW acute spinal cord injury; nervous system tumour; multiple sclerosis;
KW neurotoxin exposure; metabolic disease; diabetes; renal dysfunction;
KW eye disorder.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Modified-site 122

FT /note= "glycosylated residue"

PN W0200001815-A2.

XX 13-JAN-2000.

XX 05-JUL-1999; 99WO-DK00384.

XX 06-JUL-1998; 98DK-0000504.

XX 09-JUL-1998; 98US-0092229.

XX 19-AUG-1998; 98DK-0001048.

PR 25-AUG-1998; 98US-0097774.
PR 06-OCT-1998; 98DK-0001265.
PR 13-OCT-1998; 98US-0103908.
PR 02-JUL-1999; 99US-0347613.
XX (NEUR-) NEUROSPARCH AS.

PA Johansen TE, Blom N, Hansen C;

XX WPI; 2000-171013/15.

XX New isolated polypeptides, used for treating e.g. neurodegenerative disease or disorder, neuronal damage or neuronal disorder of the peripheral nervous system, the medulla or the spinal cord -

PS Claim 14; Page 98; 106pp; English.

XX The present sequence represents a variant of a neurotrophic factor designated neublastin. Neublastin is a member of the glial cell line-derived neurotrophic factor sub-class of the transforming growth factor-beta superfamily of neurotrophic factors. Neublastin exhibits high affinity for the GFR-alpha3-RET receptor complex. The polypeptides can be used for treating a neurodegenerative disease or disorder, cerebral ischemic neuronal damage, traumatic brain injury, peripheral neuropathy, Alzheimer's disease, Huntington's disease, Parkinson's disease, Parkinson-Plus syndromes, progressive Supranuclear Palsy, Olivopontocerebellar atrophy, Shy-Drager Syndrome, Guamanian parkinsonism dementia complex, amyotrophic lateral sclerosis, memory impairment, or a neuronal disorder of the peripheral nervous system, the medulla or the spinal cord. They can also be used for treating various neuropathies. They can also be used for treating ischemic stroke, acute brain injury, acute spinal cord injury, nervous system tumours, multiple sclerosis, exposure to neurotoxins, metabolic diseases such as diabetes or renal dysfunctions and damage caused by infectious agents, or various disorders in the eye.

XX SQ Sequence 140 AA;

Query Match 100.0%; Score 601; DB 21; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.5e-56;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGPGSARAAGARGCLRLSOLVPRALGHRSDRLVRFRCSCGRRARSPHDLSLAS 60
DB 28 agpgsaraagargclrlsqvpralglhrsdrlvrfrcsgrrarsphdlslas 87

QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113

DB 88 llgagalrpppgsrpvspccrptryeavsfmdvnsrtvdrisatagclg 140

RESULT 8

AY44774

ID AAY44774 standard; Protein; 159 AA.

XX AAY44774;

XX 17-MAY-2000 (first entry)

XX Protein-2 comprising mature human Enovin sequence.

XX Enovin; EVN; neurotrophic growth factor; chromosome lp31.3-32;
KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;
KW GDNF family receptor alpha-3; GFR alpha 3; nootropic; analgesic;
KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;
KW antidiarrhoeal; laxative; antiemetic; neurological disorder; Parkinson's;
KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;
KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;
KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;
KW Crohn's disease; bowel hypersensitivity.

XX Homo sapiens.

OS

XX FH . Location/Qualifiers
 FT Peptide 6..46
 FT /label= Pro-sequence
 FT /note= Partial pro sequence of enovin"
 FT Protein 47..159
 FT /label= Mature_Enovin
 FT /note= "Homologous to GDNF, Neurturin and Persephin"
 FT Modified-site 141..143
 FT /note= "Asn is N-glycosylated"
 FT Misc-difference 62
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 89
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 93
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 126
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 127
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 135
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 157
 FT /note= "Conserved residue characteristic of Transforming Growth Factor-beta (TGF-beta) family"
 XX WO200004050-A2.
 XX 27-JAN-2000.
 XX 14-JUL-1999; 99WO-EP05031.
 XX 14-JUL-1998; 98GB-0015283.
 XX 12-FEB-1999; 99US-0248772.
 XX 08-JUN-1999; 99US-0327668.
 XX (JANC) JANSSEN PHARM NV.
 XX Geerts HA, Masure SLJ, Meert TF, Clik M, Ver Donck LAL;
 XX WPI; 2000-182404/16.
 XX N-PSDB; AA250090.
 XX Novel human neurotrophic growth factor designated enovin used to treat neurological disorders, neuronal disorders, peripheral neuropathy, brain injury, nervous system disorders, prion associated and gastrointestinal diseases -
 XX Disclosure; Fig 3; 125pp; English.
 XX The present sequence is protein-2, comprising of partial pro sequence and mature human enovin sequence. Enovin (EVN) is a neurotrophic growth factor, that belongs to glial cell-line derived neurotrophic factor (GDNF) family. It binds to GDNF family receptor alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32. It is predominantly expressed in heart, skeletal muscle, pancreas and prostate. It has nootropic, analgesic, neuroprotective, antirheumatic, cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal, laxative and antileptic activity. It can be used to treat neurological disorders like Parkinson's, Alzheimer's and Huntington's disease, neuropathy, multiple sclerosis, endocrine neoplasia, prion associated diseases, stroke, pain, rheumatic/inflammatory diseases and gastrointestinal disorders like dyspepsia, constipation, intestinal atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
 XX Sequence 159 AA;

Query Match 100.0%; Score 601; DB 21; Length 159;
 Best Local Similarity 100.0%; Pred. No. 1.8e-56;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGGPGSRARAAGRCRLRSQLVPRALGLGHRSDLVRRFCSCGSRARSPhdLSLAS 60
 Db 47 aggpssaraagagcrlrsqvlpralglhrsdvrrfcsgscrrarsphdlsas 106
 QY 61 LIGAGALRPPGSRPVSPQCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 Db 107 llgagalrppgsrpvsqccrptryeavsfmdvnsrtvdrtsatacgcglg 159
 RESULT 9
 AAY84583
 ID AAY84583 standard; Protein; 220 AA.
 AC AAY84583;
 DT 25-JUL-2000 (first entry)
 DE Amino acid sequence of a human pre-pro-artemin polypeptide.
 KW Human; artemin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;
 KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.
 OS Homo sapiens.
 XX WO200018799-Al.
 XX 06-APR-2000.
 XX 29-SEP-1999; 99WO-US22604.
 XX 29-SEP-1998; 98US-0163283.
 XX 12-NOV-1998; 98US-0108148.
 XX 22-DEC-1998; 98US-0218698.
 XX (UNIW) UNIV WASHINGTON.
 XX Milbrandt JD, Baloh RH;
 XX WPI; 2000-293109/25.
 XX N-PSDB; AAA12540.
 XX Isolated artemin growth factor proteins and the nucleic acids that encode them, useful for treating a range of degenerative neuronal disorders such as Parkinson's disease and Huntington's disease -
 XX Claim 5; Fig 1B; 96pp; English.
 XX The present sequence represents a pre-pro-artemin growth factor protein.
 XX Artemin is a neurotrophic factor that belongs to the GDNF (glial cell line-derived neurotrophic factor)/neurturin/persephin family of growth factors and promotes differentiation, maintains mature phenotype and provides trophic support, promoting growth and survival of neurons.
 XX Artemin promotes the survival of trigeminal ganglion neurons, nodose ganglion neurons, superior cervical ganglion neurons and tyrosine-hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin is the only member of the GDNF family that binds to GFR-alpha (growth factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-tyrosine kinase) receptor complex and additionally, like GDNF and neurturin, artemin also binds to and activates GFRalpha/RET. Artemin polypeptides and polynucleotides are administered to treat peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, acute brain

CC injury, acute spinal cord injury, a nervous system tumour (e.g.
CC blastomas), multiple sclerosis, infection or enteric disease (e.g.
CC idiopathic constipation or constipation associated with Parkinson's
CC disease, spinal cord injury or use of opiate pain killers). They may
CC also be used to treat a patient suffering from small cell lung carcinoma.
XX
XX
SQ Sequence 220 AA;

Query Match 100.0%; Score 601; DB 21; Length 220;
Best Local Similarity 100.0%; Pred. No. 2.5e-56;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGPGSRRARAGARGCRLRSQVPRALGHRSDVLRFCSCGRRARSPHDLSLAS 60
Db 108 agpgsrraraagargcrlrsqvlpralghrsdelvrfcsgcrrarsphdlsias 167
QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
Db 168 llgagairpppgsrpvspccrptryeavsfmdvnsrtwtvdrilsatacgclg 220

RESULT 10
AA44776
ID AAY44776 standard; Protein; 220 AA.
XX
AC AAY44776;
XX
DT 17-MAY-2000 (first entry)
XX
DE Short splice variant of human Enovin.
XX
KW Enovin; EVN; neurotrophic growth factor; chromosome 1p31.3-32;
KW glial cell-line derived neurotrophic factor; GDNF; neuroprotective;
KW GDNF family receptor alpha-3; GFR alpha 3; nontropic; analgesic;
KW antirheumatic; cerebroprotective; antiparkinsonian; antiinflammatory;
KW antidiarrhoeal; laxative; antileptic; neurologically disorder; Parkinson's;
KW Alzheimer's; Huntington's; neuropathy; multiple sclerosis; stroke; pain;
KW endocrine neoplasia; prion; rheumatic; inflammation; gastrointestinal;
KW dyspepsia; constipation; intestinal atony; emesis; diarrhoea;
KW Crohn's disease; bowel hypersensitivity; gene therapy; splice variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..39
FT /label= Signal_Peptide
FT Peptide 40..107
FT /label= Pro-sequence
FT Misc-difference 82..220
FT /note= "This region has been claimed specifically"
FT Protein 108..220
FT /label= Mature_Enovin
FT /note= "Homologous to GDNF, Neurturin and Persephin"
FT Modified-site 202..204
FT /note= "Asn is N-glycosylated"
FT Misc-difference 123
FT /note= "Conserved residue characteristic of Transforming
FT Growth Factor-beta (TGF-beta) family"
FT Misc-difference 150
FT /note= "Conserved residue characteristic of Transforming
FT Growth Factor-beta (TGF-beta) family"
FT Misc-difference 154
FT /note= "Conserved residue characteristic of Transforming
FT Growth Factor-beta (TGF-beta) family"
FT Misc-difference 187
FT /note= "Conserved residue characteristic of Transforming
FT Growth Factor-beta (TGF-beta) family"
FT Misc-difference 188
FT /note= "Conserved residue characteristic of Transforming
FT Growth Factor-beta (TGF-beta) family"
FT Misc-difference 216
FT /note= "Conserved residue characteristic of Transforming

FT Misc-difference 218
FT /note= "Conserved residue characteristic of Transforming
FT Growth Factor-beta (TGF-beta) family"
XX
FN WO200004050-A2.
XX
XX 27-JAN-2000.
XX 14-JUL-1999; 99WO-EP05031.
XX 14-JUL-1998; 98GB-0015283.
PR 12-FEB-1999; 99US-0248772.
PR 08-JUN-1999; 99US-0327668.
XX
PA (JANC) JANSSEN PHARM NV.
XX
PI Geerts HA, Masure SLJ, Meert TP, Clk M, Ver Donck LAL;
XX WPI; 2000-182404/16.
DR N-PSDB; AAZ50091.
XX
PT Novel human neurotrophic growth factor designated enovin used to treat
PT neurological disorders, neuronal disorders, peripheral neuropathy,
PT brain injury, nervous system disorders, prion associated and
PT gastrointestinal diseases -
XX
XX Claim 11; Fig 24; 125pp; English.
XX
CC The present sequence is a short splice variant of human Enovin (EVN). EVN
CC is a neurotrophic growth factor, that belongs to glial cell-line derived
CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor
CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.
CC It is predominantly expressed in heart, skeletal muscle, pancreas and
CC prostate. It has nontropic, analgesic, neuroprotective, antirheumatic,
CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,
CC laxative and antileptic activity. It can be used to treat neurological
CC disorders like Parkinson's, Alzheimer's and Huntington's disease,
CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated
CC diseases, stroke, pain, rheumatic/inflammatory diseases and
CC gastrointestinal disorders like dyspepsia, constipation, intestinal
CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
CC EVN polynucleotide can be used in gene therapy.
XX
SQ Sequence 220 AA;
Query Match 100.0%; Score 601; DB 21; Length 220;
Best Local Similarity 100.0%; Pred. No. 2.5e-56;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGPGSRRARAGARGCRLRSQVPRALGHRSDVLRFCSCGRRARSPHDLSLAS 60
Db 108 agpgsrraraagargcrlrsqvlpralghrsdelvrfcsgcrrarsphdlsias 167
QY 61 LLGAGALRPPPGSRPVSPCCRPTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
Db 168 llgagairpppgsrpvspccrptryeavsfmdvnsrtwtvdrilsatacgclg 220
RESULT 11
AAY68710
ID AAY68710 standard; Protein; 220 AA.
XX
AC AAY68710;
XX
DT 05-MAY-2000 (first entry)
XX
DE A human pre-pro-neublastin neurotrophic factor.
KW Neurotrophic factor; neublastin; neurodegenerative disease;
KW cerebral ischemic neuronal damage; traumatic brain injury;
KW peripheral neuropathy; Alzheimer's disease; Huntington's disease;

FT /note= "Conserved residue characteristic of Transforming
 FT Growth Factor-beta (TGF-beta) family"
 FT Misc-difference 226
 FT /note= "Conserved residue characteristic of Transforming
 FT Growth Factor-beta (TGF-beta) family"
 XX WO200004050-A2.

Nels

XX 27-JAN-2000. 99WO-EP05031.

XX 14-JUL-1999; 98GB-0015283.
 XX 12-FEB-1999; 99US-0248772.
 PR 08-JUN-1999; 99US-0327668.

XX (JANC) JANSSEN PHARM NV.

XX Geerts HA, Masure SLJ, Meert TF, Cik M, Ver Donck LAL;
 PI N-PSDB; AAZ50091.

XX WPI: 2000-182404/16.
 DR N-PSDB; AAZ50091.

XX Novel human neurotrophic growth factor designated enovin used to treat
 PT neurological disorders, neuronal disorders, peripheral neuropathy,
 PT brain injury, nervous system disorders, prion associated and
 PT gastrointestinal diseases

XX Claim 11; Fig 23; 125pp; English.

XX The present sequence is a long splice variant of human Enovin (ENV). ENV
 CC is a neurotrophic growth factor, that belongs to glial cell-line derived
 CC neurotrophic factor (GDNF) family. It binds to GDNF family receptor
 CC alpha-3 (GFR alpha 3). Enovin gene is located on chromosome 1p31.3-32.
 CC It is predominantly expressed in heart, skeletal muscle, pancreas and
 CC prostate. It has neurotrophic, analgesic, neuroprotective, antirheumatic,
 CC cerebroprotective, antiparkinsonian, antiinflammatory, antidiarrhoeal,
 CC laxative and antileptic activity. It can be used to treat neurological
 CC disorders like Parkinson's, Alzheimer's and Huntington's disease.
 CC neuropathy, multiple sclerosis, endocrine neoplasia, prion associated
 CC diseases, stroke, pain, rheumatic/inflammatory diseases and
 CC gastrointestinal disorders like dyspepsia, constipation, intestinal
 CC atony, emesis, diarrhoea, Crohn's disease and bowel hypersensitivity.
 CC ENV polynucleotide can be used in gene therapy.

XX Sequence 228 AA;

Query Match 100.0%; Score 601; DB 21; Length 228;
 Best Local Similarity 100.0%; Pred. No. 2.6e-56;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRARAAGCRLRSQVPRALGLGHRSDLVFRFCGSCRRARSPHDLISLAS 60
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 LLGAGALRPPPGSRPVSPCCRRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 176 LLGAGALRPPPGSRPVSPCCRRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 228
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 15
 ID AAY84582
 AC AAY84582 standard; Protein; 229 AA.

XX AAY84582;

XX 25-JUL-2000 (first entry)

DE Amino acid sequence of a human artemin polypeptide.

XX Human; artemin; growth factor; neurotrophic factor; trophic support;
 KW neuron; trigeminal ganglion neuron; nodose ganglion neuron;

KW superior cervical ganglion neuron; midbrain neuron; Alzheimer's disease;
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
 KW Parkinson's disease; Huntington's disease; acute brain injury;
 KW acute spinal cord injury; nervous system tumour; blastoma;
 KW multiple sclerosis; infection; enteric disease; idiopathic constipation;
 KW Parkinson's disease; small cell lung carcinoma.

OS Homo sapiens.

XX WO200018799-A1.

XX 06-APR-2000.

XX 29-SEP-1999; 99WO-UP22604.

XX 29-SEP-1998; 98US-0163283.

XX 12-NOV-1998; 98US-0108148.

XX 22-DEC-1998; 98US-0218698.

XX (UNIW) UNIV WASHINGTON.

XX Milbrandt JD, Baloh RH;
 PI N-PSDB; AAA12539.

XX WPI: 2000-293109/25.
 DR N-PSDB; AAA12539.

XX Isolated artemin growth factor proteins and the nucleic acids that
 PT encode them, useful for treating a range of degenerative neuronal
 PT disorders such as Parkinson's disease and Huntington's disease -
 XX Disclosure; Fig 1A; 96pp; English.

XX The present sequence represents a human artemin growth factor protein.
 CC Artemin is a neurotrophic factor that belongs to the GDNF (glial cell
 CC line-derived neurotrophic factor)/neurturin/persephin family of growth
 CC factors and promotes differentiation, maintains mature phenotype and
 CC provides trophic support, promoting growth and survival of neurons.
 CC Artemin promotes the survival of trigeminal ganglion neurons, nodose
 CC ganglion neurons, superior cervical ganglion neurons and tyrosine-
 CC hydroxylase-expressing dopaminergic ventral midbrain neurons. Artemin
 CC is the only member of the GDNF family that binds to GFR-alpha (growth
 CC factor receptor-alpha) and activates the GFR-alpha3/RET (Ret protein-
 CC tyrosine kinase) receptor complex and additionally, like GDNF and
 CC neurturin, artemin also binds to and activates GFRalpha/RET. Artemin
 CC polypeptides and polynucleotides are administered to treat peripheral
 CC neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, ischemic stroke, acute brain
 CC injury, acute spinal cord injury, a nervous system tumour (e.g.
 CC blastomas), multiple sclerosis, infection or enteric disease (e.g.
 CC idiopathic constipation or constipation associated with Parkinson's
 CC disease, spinal cord injury or use of opiate pain killers). They may
 CC also be used to treat a patient suffering from small cell lung carcinoma.

XX Sequence 229 AA;

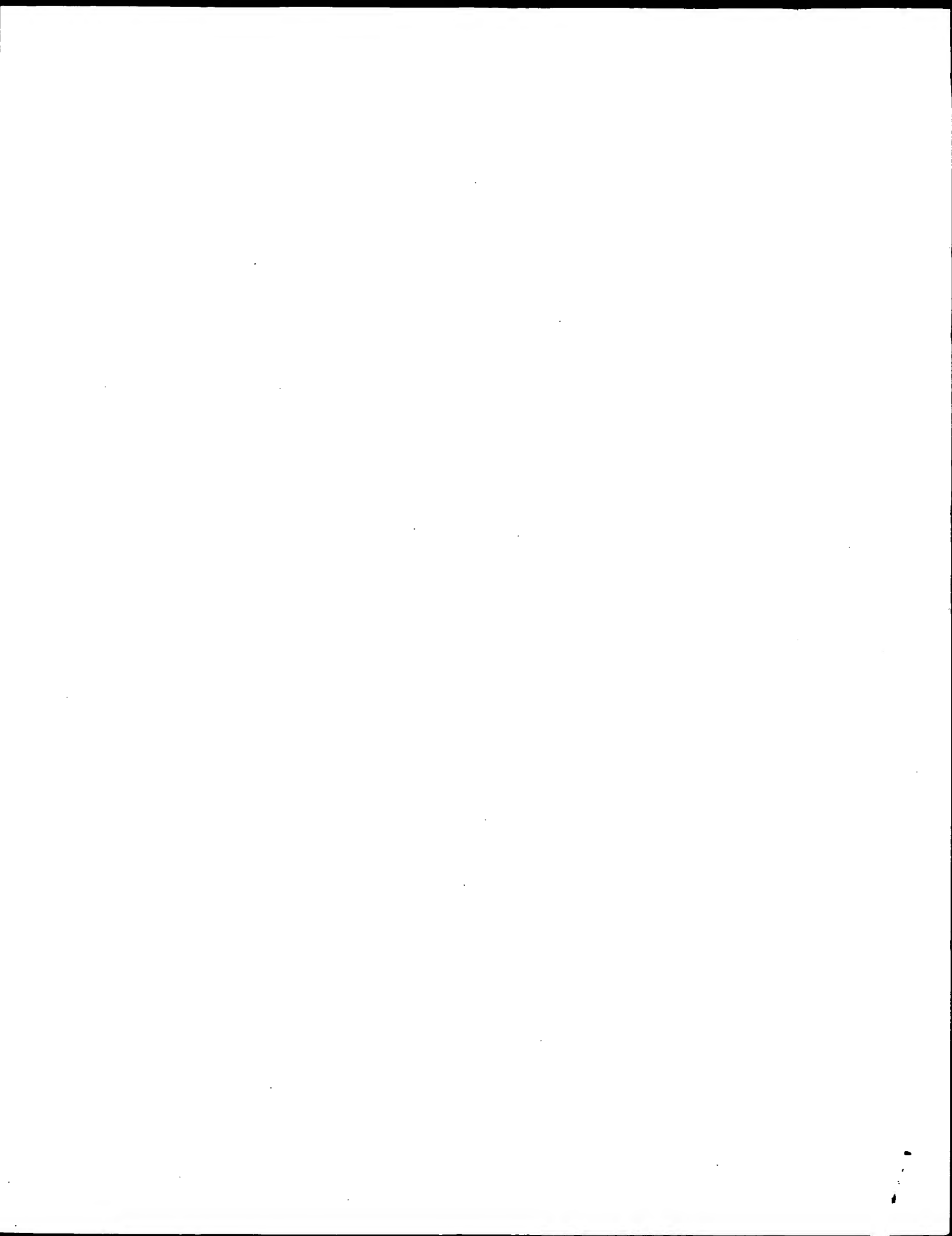
Query Match 100.0%; Score 601; DB 21; Length 229;
 Best Local Similarity 100.0%; Pred. No. 2.6e-56;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGPGSRARAAGCRLRSQVPRALGLGHRSDLVFRFCGSCRRARSPHDLISLAS 60
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 LLGAGALRPPPGSRPVSPCCRRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 113
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 176 LLGAGALRPPPGSRPVSPCCRRTRYEAVSFMDVNSTWRTVDRLSATACGCLG 215
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: August 16, 2001, 15:39:11
 Job time: 41 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:45:31 ; Search time 18.89 Seconds
(without alignments)
151.512 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 139

Sequence: 1 PPQSPRAPPAPPSPALPR.....VNSTWTVDRSLATACCLG 139

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/ptodata/2/1aa/5A-COMB.pep:*
2: /cgn2.6/ptodata/2/1aa/5B-COMB.pep:*
3: /cgn2.6/ptodata/2/1aa/6A-COMB.pep:*
4: /cgn2.6/ptodata/2/1aa/6B-COMB.pep:*
5: /cgn2.6/ptodata/2/1aa/PCTUS-COMB.pep:*
6: /cgn2.6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 9 | 6.5 | 45 | 3 | US-08-236-886-2 |
| 2 | 9 | 6.5 | 429 | 1 | US-07-672-483-3 |
| 3 | 9 | 6.5 | 429 | 6 | 5182195-13 |
| 4 | 8 | 5.8 | 89 | 4 | US-08-931-858E-223 |
| 5 | 8 | 5.8 | 96 | 4 | US-08-931-858E-221 |
| 6 | 8 | 5.8 | 133 | 4 | US-08-931-858E-132 |
| 7 | 8 | 5.8 | 156 | 4 | US-08-931-858E-217 |
| 8 | 8 | 5.8 | 196 | 2 | US-08-481-814A-9 |
| 9 | 8 | 5.8 | 365 | 2 | US-08-481-814A-10 |
| 10 | 8 | 5.8 | 437 | 1 | US-08-136-119-4 |
| 11 | 8 | 5.8 | 437 | 1 | US-07-882-711-2 |
| 12 | 8 | 5.8 | 437 | 2 | US-08-723-415B-13 |
| 13 | 8 | 5.8 | 437 | 2 | US-08-481-814A-6 |
| 14 | 8 | 5.8 | 437 | 2 | US-08-462-174-2 |
| 15 | 8 | 5.8 | 437 | 3 | US-08-801-092-1 |
| 16 | 8 | 5.8 | 437 | 4 | US-09-189-627A-13 |
| 17 | 8 | 5.8 | 476 | 2 | US-08-139-937-14 |
| 18 | 8 | 5.8 | 476 | 5 | PCT-US93-11310-14 |
| 19 | 7 | 5.0 | 7 | 4 | US-09-173-941-15 |
| 20 | 7 | 5.0 | 13 | 4 | US-09-258-754-108 |
| 21 | 7 | 5.0 | 13 | 4 | US-09-042-107-108 |
| 22 | 7 | 5.0 | 22 | 4 | US-08-602-999A-448 |
| 23 | 7 | 5.0 | 28 | 3 | US-08-256-747C-43 |
| 24 | 7 | 5.0 | 28 | 4 | US-08-834-130A-43 |
| 25 | 7 | 5.0 | 94 | 1 | US-08-519-777-31 |
| 26 | 7 | 5.0 | 94 | 1 | US-08-742-035-31 |
| 27 | 7 | 5.0 | 94 | 2 | US-08-777-019-31 |

28 7 5.0 94 2 US-08-777-143-31 Sequence 31, Appl
29 7 5.0 94 3 US-08-775-414-31 Sequence 31, Appl
30 7 5.0 94 4 US-08-931-858E-31 Sequence 31, Appl
31 7 5.0 94 4 US-08-981-739-31 Sequence 31, Appl
32 7 5.0 95 4 US-08-931-858E-173 Sequence 173, App
33 7 5.0 95 4 US-08-981-739-173 Sequence 173, App
34 7 5.0 97 1 US-08-443-568B-18 Sequence 18, Appl
35 7 5.0 97 5 PCT-US94-06997-18 Sequence 18, Appl
36 7 5.0 102 1 US-08-519-777-1 Sequence 1, Appl
37 7 5.0 102 1 US-08-742-035-1 Sequence 1, Appl
38 7 5.0 102 2 US-08-777-019-1 Sequence 1, Appl
39 7 5.0 102 2 US-08-777-143-1 Sequence 1, Appl
40 7 5.0 102 3 US-09-106-486-1 Sequence 1, Appl
41 7 5.0 102 3 US-08-775-414-1 Sequence 1, Appl
42 7 5.0 102 4 US-08-931-858E-1 Sequence 1, Appl
43 7 5.0 102 4 US-08-981-739-1 Sequence 1, Appl
44 7 5.0 103 3 US-09-106-486-4 Sequence 4, Appl
45 7 5.0 104 3 US-08-775-414-79 Sequence 79, Appl

ALIGNMENTS

RESULT 1
US-08-236-886-2
; Sequence 2, Application US/08236886
; Patent No. 6110662
; GENERAL INFORMATION:
; APPLICANT: Fong, Steven K.
; APPLICANT: Hadlock, Kenneth G.
; APPLICANT: Chow, Theresa P.
; TITLE OF INVENTION: HTLV-I/HTLV-II Assay and Method
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Peter Dehlinger
; STREET: 350 Cambridge Avenue, Suite 300
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,886
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/840,906
; FILING DATE: 24-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4500-0097
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Figure 3, p19-R45, p19-C45
US-08-236-886-2

Query Match 6.5%; Score 9; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.26;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
| | | | | | | |
Db 10 PSRPAPPPP 18

RESULT 2

US-07-672-483-3
; Sequence 3, Application US/07672483
; Patent No. 5359029
; GENERAL INFORMATION:
; APPLICANT: LACROIX, Martial
; APPLICANT: ZREIN, Maan
; TITLE OF INVENTION: PEPTIDES AND ANALOGUES AND MIXTURES
; TITLE OF INVENTION: THEREOF FOR DETECTING ANTIBODIES TO HTLV-I AND HTLV-II
; TITLE OF INVENTION: VIRUSES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; CITY: 875 Third Avenue
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/672.483
; FILING DATE: 19910302
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/554,258
; FILING DATE: 18-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY JR, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: IAF CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0674
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 429 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-672-483-3

Query Match 6.5%; Score 9; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
| | | | | | | |
Db 95 PSRPAPPPP 103

RESULT 3

5182195-13
; Patent No. 5182195
; APPLICANT: NAKAHAMA, KAZUO; KATISHO, YOSHITHIKO; YOSHIMURA, KOJI
; TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE
; DEFICIENT YEASTS
; NUMBER OF SEQUENCES: 71
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/269,140

; FILING DATE: 09-NOV-1988
; SEQ ID NO:13:
; LENGTH: 429
5182195-13

Query Match 6.5%; Score 9; DB 6; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
| | | | | | | |
Db 95 PSRPAPPPP 103

RESULT 4

US-08-931-858E-223
; Sequence 223, Application US/08931858E
; Patent No. 6222022
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, EUGENE M
; APPLICANT: MILBRANDT, JEFFREY D
; APPLICANT: KOTZBAUER, PAUL T
; APPLICANT: LAMPE, PATRICIA A
; APPLICANT: KLEIN, ROBERT
; APPLICANT: DESAUVAGE, FRED
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,858E
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 971486
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-6092
; TELEFAX: 314-727-5188
; INFORMATION FOR SEQ ID NO: 223:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-931-858E-223

Query Match 5.8%; Score 8; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PCCRPTRY.112
| | | | | | | |
Db 57 PCCRPTRY 64

RESULT 5

US-08-931-858E-221
; Sequence 221, Application US/08931858E

Patent No. 6222022
 GENERAL INFORMATION:
 APPLICANT: JOHNSON, EUGENE M
 APPLICANT: MILBRANDT, JEFFREY D
 APPLICANT: KOTZBAUER, PAUL T
 APPLICANT: LAMPE, PATRICIA A
 APPLICANT: KLEIN, ROBERT
 APPLICANT: DESAUVAGE, FRED
 TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
 NUMBER OF SEQUENCES: 239
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: HOWELL & HAFERKAMP, L.C.
 STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
 CITY: ST. LOUIS
 STATE: MO
 COUNTRY: USA
 ZIP: 63105
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/931,858E
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: HOLLAND, DONALD R.
 REGISTRATION NUMBER: 35,197
 REFERENCE/DOCKET NUMBER: 971486
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 314-727-5188
 TELEFAX: 314-727-6092
 INFORMATION FOR SEQ ID NO: 221:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 96 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-931-858E-221

Query Match 5.8%; Score 8; DB 4; Length 96;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PCCRPTRY 112
 |||||
 DB 62 PCCRPTRY 69

RESULT 6
 US-08-931-858E-132
 Sequence 132, Application US/08931858E
 Patent No. 6222022
 GENERAL INFORMATION:
 APPLICANT: JOHNSON, EUGENE M
 APPLICANT: MILBRANDT, JEFFREY D
 APPLICANT: KOTZBAUER, PAUL T
 APPLICANT: LAMPE, PATRICIA A
 APPLICANT: KLEIN, ROBERT
 APPLICANT: DESAUVAGE, FRED
 TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
 NUMBER OF SEQUENCES: 239
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: HOWELL & HAFERKAMP, L.C.
 STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
 CITY: ST. LOUIS
 STATE: MO
 COUNTRY: USA
 ZIP: 63105
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/931,858E
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: HOLLAND, DONALD R.
 REGISTRATION NUMBER: 35,197
 REFERENCE/DOCKET NUMBER: 971486
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 314-727-5188
 TELEFAX: 314-727-6092
 INFORMATION FOR SEQ ID NO: 132:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 133 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-931-858E-132

Query Match 5.8%; Score 8; DB 4; Length 133;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PCCRPTRY 112
 |||||
 DB 99 PCCRPTRY 106

RESULT 7
 US-08-931-858E-217
 Sequence 217, Application US/08931858E
 Patent No. 6222022
 GENERAL INFORMATION:
 APPLICANT: JOHNSON, EUGENE M
 APPLICANT: MILBRANDT, JEFFREY D
 APPLICANT: KOTZBAUER, PAUL T
 APPLICANT: LAMPE, PATRICIA A
 APPLICANT: KLEIN, ROBERT
 APPLICANT: DESAUVAGE, FRED
 TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR
 NUMBER OF SEQUENCES: 239
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: HOWELL & HAFERKAMP, L.C.
 STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
 CITY: ST. LOUIS
 STATE: MO
 COUNTRY: USA
 ZIP: 63105
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/931,858E
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: HOLLAND, DONALD R.
 REGISTRATION NUMBER: 35,197
 REFERENCE/DOCKET NUMBER: 971486
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 314-727-5188
 TELEFAX: 314-727-6092
 INFORMATION FOR SEQ ID NO: 217:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 156 amino acids

```
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-931-858E-217

Query Match 5.8%; Score 8; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PCCRPTX 112
Db 122 PCCRPTX 129

RESULT 8
US-08-481-814A-9
; Sequence 9, Application US/08481814A
; Patent No. 5869040
; GENERAL INFORMATION:
; APPLICANT: Hsu, Yen-Ming
; TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan, Warren A
; REFERENCE/DOCKET NUMBER: A001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-679-2000
; TELEFAX: 617-679-2838
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 196 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-814A-9

Query Match 5.8%; Score 8; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LLGAGALR 94
Db 19 LLGAGALR 26

RESULT 9
US-08-481-814A-10
; Sequence 10, Application US/08481814A
; Patent No. 5869040
; GENERAL INFORMATION:
; APPLICANT: Hsu, Yen-Ming
```

```
; TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,814A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan, Warren A
; REFERENCE/DOCKET NUMBER: A001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-679-2000
; TELEFAX: 617-679-2838
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-814A-10

Query Match 5.8%; Score 8; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LLGAGALR 94
Db 19 LLGAGALR 26

RESULT 10
US-08-136-119-4
; Sequence 4, Application US/08136119
; Patent No. 5473056
; GENERAL INFORMATION:
; APPLICANT: Heimbrosk, David C.
; APPLICANT: Hoyle, Mona I.
; APPLICANT: Oliff, Allen I.
; TITLE OF INVENTION: E 2F-2, A NOVEL MAMMALIAN TRANSCRIPTION
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Muthard
; STREET: P.O. Box 2000, 126 Lincoln Avenue
; CITY: Rahway
; STATE: N.J.
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,119
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

NAME: Muthard, David A.
REGISTRATION NUMBER: 35,297
REFERENCE/DOCKET NUMBER: 19042
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-3903
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-136-119-4

Query Match 5.8%; Score 8; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94
Db 19 LLGAGALR 26

RESULT 11
US-07-882-711-2
Sequence 2, Application US/07882711
Patent No. 5759803
GENERAL INFORMATION:
APPLICANT: Kaelin Jr., William G.
APPLICANT: Flemington, Erik
APPLICANT: Sellers, William
APPLICANT: Decaprio, James A.
APPLICANT: Livingston, David M.
TITLE OF INVENTION: Retinoblastoma Associated Protein-1 cDNA
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Choate, Hall & Stewart
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/882,711
FILING DATE: 19920513
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kennedy, Bill
REGISTRATION NUMBER: 33,407
REFERENCE/DOCKET NUMBER: DFCI#236
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-5020
TELEFAX: (617) 227-7566
TELEX: 289374
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-07-882-711-2

Query Match 5.8%; Score 8; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 87 LLGAGALR 94
Db 19 LLGAGALR 26

RESULT 12
US-08-723-415B-13
Sequence 13, Application US/08723415B
Patent No. 5859199
GENERAL INFORMATION:
APPLICANT: Lathangue, Nicholas B.
APPLICANT: delaluna, Susana
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 No. 5859199th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,415B
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9610195.1
FILING DATE: 15-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 117-220
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-723-415B-13

Query Match 5.8%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94
Db 19 LLGAGALR 26

RESULT 13
US-08-481-814A-6
Sequence 6, Application US/08481814A
Patent No. 5869040
GENERAL INFORMATION:
APPLICANT: Hsu, Yen-Ming
TITLE OF INVENTION: GENE THERAPY METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Biogen, Inc.
STREET: 14 Cambridge Center
CITY: Cambridge

STATE: Massachusetts
COUNTRY: United States of America
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,814A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan, Warren A
REFERENCE/DOCKET NUMBER: A001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-679-2000
TELEFAX: 617-679-2838
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-814A-6

Query Match 5.8%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94
|||||||
Db 19 LLGAGALR 26

RESULT 14
US-08-462-174-2
Sequence 2, Application US/08462174
Patent No. 5981723
GENERAL INFORMATION:
APPLICANT: Kaelin Jr., William G.
APPLICANT: Flemington, Erik
APPLICANT: Sellers, William
APPLICANT: DeCaprio, James A.
APPLICANT: Livingston, David M.
TITLE OF INVENTION: Retinoblastoma Associated Protein-1 cDNA
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Choate, Hall & Stewart
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,174
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/882,711
FILING DATE: 13-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kennedy, Bill
REGISTRATION NUMBER: 33,407

REFERENCE/DOCKET NUMBER: DFCI#236
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-5020
TELEFAX: (617) 227-7566
TELEX: 289374
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-462-174-2

Query Match 5.8%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94
|||||||
Db 19 LLGAGALR 26

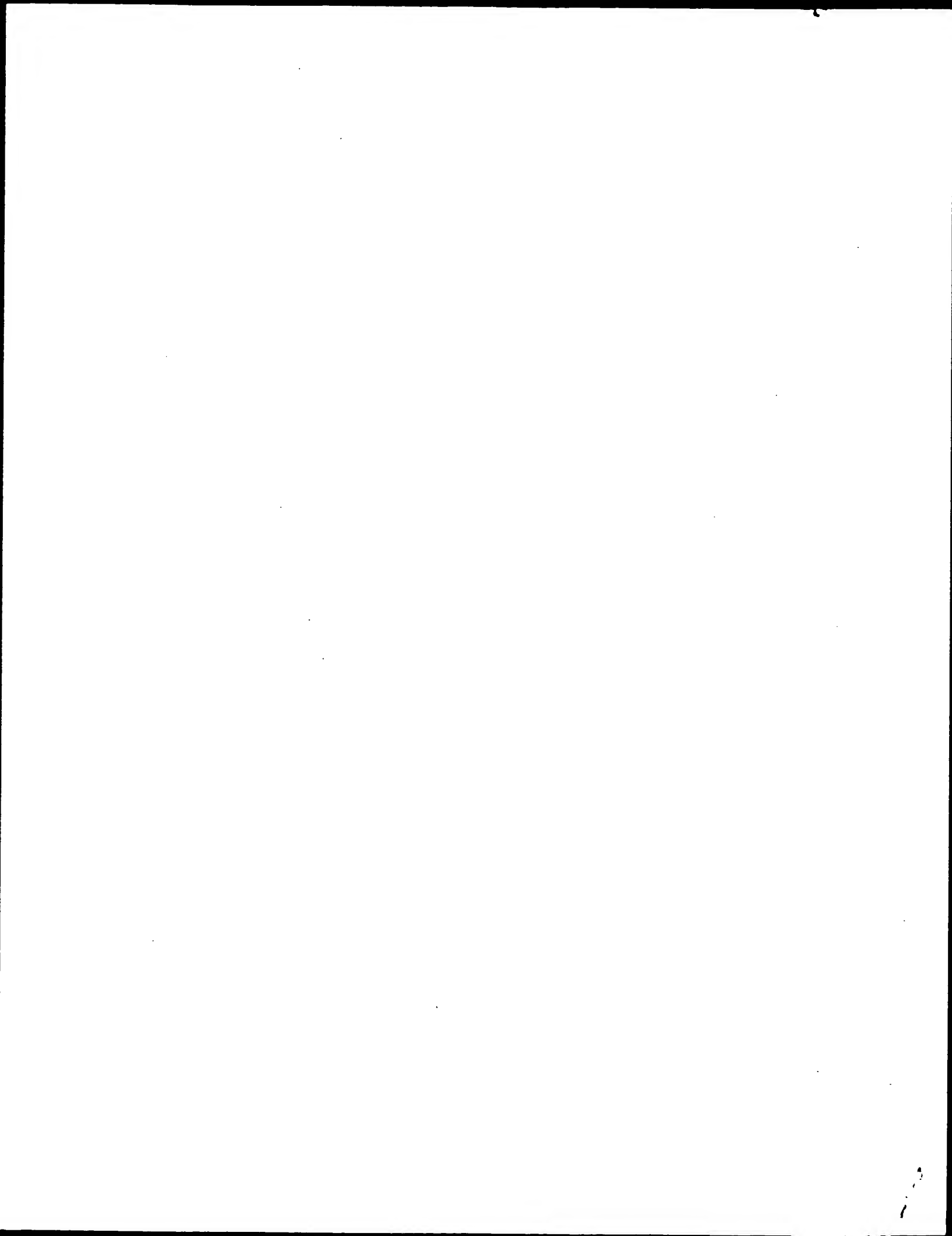
RESULT 15
US-08-801-092-1
Sequence 1, Application US/08801092
Patent No. 6074850
GENERAL INFORMATION:
APPLICANT: Antelman, Douglas
APPLICANT: Gregory, Richard J.
APPLICANT: Wills, Kenneth N.
TITLE OF INVENTION: Tissue Specific Expression of
TITLE OF INVENTION: Retinoblastoma Protein
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,092
FILING DATE: 14-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,517
FILING DATE: 15-NOV-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 016930-001020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 703-576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-801-092-1

Query Match 5.8%; Score 8; DB 3; Length 437;

Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LLGAGALR 94
Db 19 LLGAGALR 26

Search completed: August 16, 2001, 15:45:32
Job time: 282 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2001, 15:45:06 ; Search time 21.79 Seconds
(without alignments)
485,923 Million cell updates/sec

Title: US-09-357-349-4
Perfect score: 139
Sequence: 1 PPOPSRPPPPPPPPSALPR.....VNSTWRTVDRLSATACGLG 139

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 9 | 6.5 | 113 | 2 T00563 | hypothetical prote |
| 2 | 9 | 6.5 | 248 | 2 B86044 | espf [imported] - |
| 3 | 9 | 6.5 | 429 | 1 FOLJGH | probable gag poly |
| 4 | 9 | 6.5 | 429 | 1 FOLJCN | gag polypeptide - |
| 5 | 9 | 6.5 | 429 | 2 S06073 | gag polypeptide - |
| 6 | 9 | 6.5 | 814 | 2 G02390 | disintegrin-like m |
| 7 | 8 | 5.8 | 75 | 2 T26916 | NADH dehydrogenase |
| 8 | 8 | 5.8 | 108 | 2 S28241 | hypothetical prote |
| 9 | 8 | 5.8 | 274 | 2 T04619 | hypothetical prote |
| 10 | 8 | 5.8 | 393 | 2 T32533 | hypothetical prote |
| 11 | 8 | 5.8 | 430 | 2 A56209 | transcription fact |
| 12 | 8 | 5.8 | 437 | 2 JC4929 | transcription fact |
| 13 | 8 | 5.8 | 484 | 2 S58868 | G protein-coupled |
| 14 | 8 | 5.8 | 678 | 2 B70913 | probable penicilli |
| 15 | 8 | 5.8 | 810 | 2 C70791 | probable pona' pro |
| 16 | 8 | 5.8 | 1111 | 2 T23047 | hypothetical prote |
| 17 | 8 | 5.8 | 1127 | 2 T32404 | hypothetical prote |
| 18 | 8 | 5.8 | 1366 | 2 B84924 | hypothetical prote |
| 19 | 8 | 5.8 | 2526 | 2 T20531 | hypothetical prote |
| 20 | 8 | 5.8 | 2722 | 2 T20532 | hypothetical prote |
| 21 | 7 | 5.0 | 34 | 2 S02382 | probable membrane |
| 22 | 7 | 5.0 | 82 | 2 T30135 | hypothetical prote |
| 23 | 7 | 5.0 | 107 | 2 E72710 | hypothetical prote |
| 24 | 7 | 5.0 | 133 | 2 E70507 | hypothetical prote |
| 25 | 7 | 5.0 | 133 | 2 I51960 | wt1 zinc-finger ho |
| 26 | 7 | 5.0 | 138 | 2 PC1195 | genome polypeptide |
| 27 | 7 | 5.0 | 161 | 2 JE0306 | Id4 protein - huma |
| 28 | 7 | 5.0 | 161 | 2 G01855 | helix-loop-helix p |
| 29 | 7 | 5.0 | 161 | 2 S43260 | helix-loop-helix p |

ALIGNMENTS

RESULT 1

T00563

hypothetical protein At2g39530 [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein F12L6.19

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 02-Mar-2001

C:Accession: T00563; E84818

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K

submitted to the EMBL Data Library, July 1998

A:Description: Arabidopsis thaliana chromosome II BAC F12L6 genomic sequence.

A:Reference number: Z14168

A:Accession: T00563

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-113 <ROU>

A:Cross-references: EMBL:AC004218; NID:g3355463; PID:g3355482

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: E84818

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-113 <STO>

A:Cross-references: GB:AE002093; NID:g3355482; PIDN:AAC27844.1; GSPDB:GN00139

C:Genetics:

A:Gene: F12L6.19; At2g39530

A:Map position: 2

A:Introns: 60/2

C:Superfamily: Arabidopsis thaliana hypothetical protein At2g39530

Query Match

Best Local Similarity 100.0%; Score 9; DB 2; Length 113;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 APPPPAPPS 16

Db 2 APPPPAPPS 10

RESULT 2

B86044

espf [imported] - Escherichia coli (strain O157:H7)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001

C:Accession: B86044

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

filler, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apoda

Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: B86044
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-248 <STO>
 A:Cross-references: GB:AE005174; NID:gl2518435; PIDN:AAG58814.1; GSPDB:GN00145; UWGP:Z51
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: espf

Query Match 6.5%; Score 9; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
 DB 73 PSRPAPPPP 81
 |||||

RESULT 3

FOLJGH
 C:Probable gag polyprotein - human T-cell lymphotropic virus type 1
 C:Species: human T-cell lymphotropic virus type 1, HTLV-1
 A:Note: host Homo sapiens (man)
 C:Date: 14-Nov-1983 #sequence_revision 27-Nov-1985 #text_change 26-Apr-1996
 C:Accession: B93954; A91315; A03943
 R:Seiki, M.; Hattori, S.; Hirayama, Y.; Yoshida, M.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3618-3622, 1983
 A:Title: Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus
 A:Reference number: A93954; MUID:83221647
 A:Accession: B93954
 A:Molecule type: DNA
 A:Residues: 1-429 <SEI>
 A:Experimental source: strain ATK
 R:Copeland, T.D.; Oroszlan, S.; Kalyanaraman, V.S.; Sarngadharan, M.G.; Gallo, R.C.
 FEBS Lett. 162, 390-395, 1983
 A:Title: Complete amino acid sequence of human T-cell leukemia virus structural protein
 A:Reference number: A91315; MUID:84029174
 A:Accession: A91315

A:Molecule type: protein
 A:Residues: 345-415, 'T', 417-429 <COP>
 C:Comment: This protein is synthesized as a gag-pol polyprotein.

C:Genetics:

A:Gene: gag

C:Superfamily: mammalian retrovirus gag polyprotein II
 C:Keywords: core protein; polyprotein
 F:1-344/Product: gag polyprotein #status predicted <GAG>
 F:345-429/Product: protein p15 #status experimental <C15>

Query Match 6.5%; Score 9; DB 1; Length 429;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
 DB 95 PSRPAPPPP 103
 |||||

RESULT 4

FOLJCN
 gag polyprotein - human T-cell lymphotropic virus type 1 (isolate Caribbean)
 N:Conatus: major core protein p19; major core protein p24; nucleic acid-binding protein
 C:Species: human T-cell lymphotropic virus type 1, HTLV-1
 A:Note: host Homo sapiens (man)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jun-2000
 C:Accession: A28136
 R:Malik, K.T.A.; Even, J.; Karpas, A.
 J. Gen. Virol. 69, 1695-1710, 1988
 A:Title: Molecular cloning and complete nucleotide sequence of an adult T cell leukaemia

bers of the ATL/HTLV-I subgroup.
 A:Reference number: A92797; MUID:88274338
 A:Accession: A28136
 A:Molecule type: DNA
 A:Residues: 1-429 <MAL>
 A:Cross-references: GB:DL3784; GB:D00294; NID:g221866; PIDN:BAA02929.1; PID:g221867
 C:Genetics:
 A:Gene: gag
 C:Superfamily: mammalian retrovirus gag polyprotein II
 C:Keywords: core protein; polyprotein
 F:1-130/Product: major core protein p19 #status predicted <P19>
 F:131-344/Product: major core protein p24 #status predicted <P24>
 F:345-429/Product: nucleic acid-binding protein p15 #status predicted <P15>

Query Match 6.5%; Score 9; DB 1; Length 429;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
 DB 95 PSRPAPPPP 103
 |||||

RESULT 5

S06073
 gag polyprotein - human T-cell lymphotropic virus type 1
 C:Species: human T-cell lymphotropic virus type 1, HTLV-1
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 26-Aug-1999
 C:Accession: S06073
 R:Gray, G.S.; Bartman, T.; White, M.
 Nucleic Acids Res. 17, 7998, 1989
 A:Title: Nucleotide sequence of the core (gag) gene from HTLV-1 isolate MT-2.
 A:Reference number: S06073; MUID:90016893
 A:Accession: S06073
 A:Molecule type: mRNA
 A:Residues: 1-429 <GRA>

A:Cross-references: EMBL:X15951; NID:g60425; PIDN:CAA34075.1; PID:g60426
 C:Superfamily: mammalian retrovirus gag polyprotein II
 C:Keywords: polyprotein

Query Match 6.5%; Score 9; DB 2; Length 429;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRPAPPPP 12
 DB 95 PSRPAPPPP 103
 |||||

RESULT 6

G02390
 disintegrin-like metalloproteinase MDC15 (EC 3.4.24.-) - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 31-Dec-2000
 C:Accession: G02390; PC4263
 R:Herren, B.; Raines, E.W.; Ross, R.
 submitted to the EMBL Data Library, January 1996

A:Reference number: H01157
 A:Accession: G02390
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-814 <HER>

A:Cross-references: EMBL:U46005; NID:g1335871; PIDN:AAC51112.1; PID:g1335872
 R:McKie, N.; Edwards, T.; Dallas, D.J.; Houghton, A.; Stringer, B.; Graham, R.; Russe
 Biochem. Biophys. Res. Commun. 230, 335-339, 1997
 A:Title: Expression of members of a novel membrane linked metalloproteinase family (A
 A:Reference number: PC4263; MUID:97168971
 A:Accession: PC4263
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-461 <MCK>

A:Experimental source: articular chondrocyte
 C:Comment: this protein is a membrane bound protein and involved in cell/cell and cell/m
 C:Superfamily: mouse meltrin alpha; disintegrin homology
 F:Keywords: hydrolase; metalloproteinase; zinc
 F:420-503/Domain: disintegrin homology <Dis>
 F:348,352,358/Binding site: zinc (His) #status predicted
 F:349/Active site: Glu #status predicted

Query Match 6.5%: Score 9; DB 2; Length 814;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSRAPPPP 12
 |||||
 DB 799 PSRAPPPP 807

RESULT 7
 T26916
 hypothetical protein Y45F10B.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T26916
 R:McMurry, A.
 submitted to the EMBL Data Library, January 1998
 A:Reference number: Z20286
 A:Accession: T26916
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-75 <WIL>
 A:Cross-references: EMBL:AL021487; PIDN:CAA16354.1; GSPDB:GN00022; CESP:Y45F10B.3
 A:Experimental source: clone Y45F10B
 C:Genetics:
 A:Gene: CESP:Y45F10B.3
 A:Map position: 4

Query Match 5.8%: Score 8; DB 2; Length 75;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PAPPPAP 14
 |||||
 DB 43 PAPPPAP 50

RESULT 8
 S28241
 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain CI-AGGG - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 22-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
 C:Accession: S28241
 R:Walker, J.E.; Arizmendi, J.M.; Dupuis, A.; Fearnley, I.M.; Finel, M.; Medd, S.M.; Pilk
 J. Mol. Biol. 226, 1051-1072, 1992
 A:Title: Sequences of 20 subunits of NADH: ubiquinone oxidoreductase from bovine heart m
 A:Reference number: S28237; MUID:92389317
 A:Accession: S28241
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-108 <NAL>
 A:Cross-references: EMBL:X63216; NID:g231; PIDN:CAA44901.1; PID:g232
 C:Keywords: NAD; oxidoreductase

Query Match 5.8%: Score 8; DB 2; Length 108;
 Best Local Similarity 100.0%; Pred. No. 2.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ARAAGARG 41
 |||||
 DB 26 ARAAGARG 33

RESULT 9
 T04619

hypothetical protein F2009.170 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999
 C:Accession: T04619
 R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.
 submitted to the Protein Sequence Database, October 1998

A:Reference number: Z15380
 A:Accession: T04619
 A:Molecule type: DNA
 A:Residues: 1-274 <BEV>
 A:Cross-references: EMBL:AL021749
 A:Experimental source: cultivar Columbia; BAC clone F2009
 C:Genetics:
 A:Map position: 4
 A:Introns: 97/3; 151/3; 210/1
 A:Note: F2009.170

Query Match 5.8%: Score 8; DB 2; Length 274;
 Best Local Similarity 100.0%; Pred. No. 6.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PAPPPAP 14
 |||||
 DB 209 PAPPPAP 216

RESULT 10
 T32533
 hypothetical protein CD4.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T32533
 R:Du, Z.; Scheet, P.
 submitted to the EMBL Data Library, December 1997
 A:Description: The sequence of C. elegans cosmid CD4.
 A:Reference number: Z21185
 A:Accession: T32533
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-393 <DUZ>
 A:Cross-references: EMBL:AF036594; PIDN:AAB88347.1; GSPDB:GN00022; CESP:CD4.9
 A:Experimental source: strain Bristol N2; clone CD4
 C:Genetics:
 A:Gene: CESP:CD4.9
 A:Map position: 4
 A:Introns: 21/2; 71/2; 100/1; 195/3; 209/3; 319/2; 379/1

Query Match 5.8%: Score 8; DB 2; Length 393;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 SRAPPPP 12
 |||||
 DB 83 SRAPPPP 90

RESULT 11
 A56209
 transcription factor E2F1 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
 C:Accession: A56209
 R:Li, Y.; Slansky, J.E.; Myers, D.J.; Drinkwater, N.R.; Kaelin, W.G.; Farnham, P.J.
 Mol. Cell. Biol. 14, 1861-1869, 1994
 A:Title: Cloning, chromosomal location, and characterization of mouse E2F1.
 A:Reference number: A56209; MUID:94158858
 A:Accession: A56209
 A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-430 <L1A>

A:Cross-references: GB:L21973; NID:g1091479; PIDN:AAA83217.1; PID:g405524

Query Match 5.8%; Score 8; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LLGAGALR 94
Db 17 LLGAGALR 24

RESULT 12

JC4929 transcription factor E2F1 - human
N:Alternate names: APl2 protein; retinoblastoma-associated protein; retinoblastoma-binding protein

C:Species: Homo sapiens (man)

C>Date: 22-Oct-1996 #sequence_revision 01-Nov-1996 #text_change 02-Jun-2000

C:Accession: JC4929; A45032; A42998; A42997; I54091

C:Neuman, E.; Sellers, W.R.; McNeill, J.A.; Lawrence, J.B.; Keelin Jr., W.G.

Gene 173, 163-169, 1996

A:Title: Structure and partial genomic sequence of the human E2F1 gene.

A:Reference number: JC4929; MUID:97082961

A:Accession: JC4929

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-437 <NEU>

A:Cross-references: GB:U47675; NID:g1594281; GB:U47676; NID:g1594282; GB:U47677; NID:g1594283

A:Experimental source: placenta

R:Shan, B.; Zhu, X.; Chen, P.L.; Durfee, T.; Yang, Y.; Sharp, D.; Lee, W.H.

Mol. Cell. Biol. 12, 5620-5631, 1992

A:Title: Molecular cloning of cellular genes encoding retinoblastoma-associated proteins

A:Reference number: A45032; MUID:93078763

A:Accession: A45032

A:Molecule type: mRNA

A:Residues: 'LCRRPGRDRAAEACRHGPPPPVTRAAVSV', 1-312, 'S', 314-321, 'N', 323-328, 'T', 329-330, 332-335, 1992

A:Cross-references: GB:S49592; NID:g260573; PIDN:AA824289.1; PID:g260574

A:Note: sequence extracted from NCBI backbone (NCBIP:119095)

A:Note: the authors are uncertain whether Met-1 is the initiator or whether translation

R:Hein, K.; Lees, J.A.; Vidal, M.; Dyson, N.; Harlow, E.; Fattaey, A.

Cell 70, 337-350, 1992

A:Title: A cDNA encoding a PRB-binding protein with properties of the transcription factor

A:Reference number: A42998; MUID:92346720

A:Accession: A42998

A:Molecule type: mRNA

A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <HEL>

A:Cross-references: GB:M96577; NID:g181917; PIDN:AAA35782.1; PID:g181918

A:Experimental source: Nalm 6 pre-B leukemia cell line

R:Kaelin Jr., W.G.; Krek, W.; Sellers, W.R.; DeCaprio, J.A.; Ajchenbaum, F.; Fuchs, C.S.

Cell 70, 351-364, 1992

A:Title: Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F

A:Reference number: A42997; MUID:92346721

A:Accession: A42997

A:Molecule type: mRNA

A:Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <KAE>

A:Cross-references: GB:U13851; NID:g595713

A:Experimental source: Akata cells; expression vector pGEX-2TK

R:Johnson, D.G.; Ohtani, K.; Nevins, J.R.

Genes Dev. 8, 1514-1525, 1994

A:Title: Autoregulatory control of E2F1 expression in response to positive and negative

A:Reference number: I54091; MUID:95047311

A:Accession: I54091

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-88, 'R', 'T', 122-123, 'TPR', 127, 'QRR', 297-299, 'PRR', 308-309, 'RA', 312, 'C' <RES>

A:Cross-references: GB:S74230; NID:g712816; PIDN:AAD14150.1; PID:g4261850

C:Genetics:

A:Gene: GDB:E2F1

A:Cross-references: GDB:I34561; OMIM:189971

A:Map position: 20q11-20q11
A:Introns: 87/3 118/1 191/2 242/2 280/3 356/1
C:Keywords: DNA binding; transcription factor
F:67-108/Region: cyclin box #status predicted
F:118-190/Domain: DNA binding #status predicted <DNA>
F:191-241/Region: 7-residue repeats

Query Match 5.8%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LLGAGALR 94
Db 19 LLGAGALR 26

RESULT 13

S58868

G protein-coupled receptor GCR1 - migratory locust

C:Species: Locusta migratoria (migratory locust)

C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 17-Nov-2000

C:Accession: S58868; S58869

R:Vanden Broeck, J.; Vulsteke, V.; Huybrechts, R.; De Loof, A.

J. Neurochem. 64, 2387-2395, 1995

A:Title: Characterization of a cloned locust tyramine receptor cDNA by functional exp

A:Reference number: S58868; MUID:95279966

A:Accession: S58868

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-484 <VAN>

A:Cross-references: EMBL:X69520; NID:g871404; PIDN:CAA49268.1; PID:g871405

A:Accession: S58869

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-307, 'D', 309-338, 'K', 340-484 <VA>

A:Cross-references: EMBL:X69521; NID:g871406; PIDN:CAA49269.1; PID:g871407

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992

C:Superfamily: octopamine receptor type I

C:Keywords: G protein-coupled receptor

Query Match 5.8%; Score 8; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PPPPAPPS 16
Db 286 PPPPAPPS 293

RESULT 14

B70913

probable penicillin-binding protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: B70913

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-678 <COL>

A:Cross-references: GB:B280775; GB:AL123456; NID:g3250715; PIDN:CAB02529.1; PID:g32507

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: ponA

C:Superfamily: penicillin-binding protein 1B

Query Match 5.8%; Score 8; DB 2; Length 678;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APPPPAPP 15

|||||||

Db 663 APPPPAPP 670

RESULT 15

C70791

Probable ponA' protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: C70791

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Reltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A: Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A: Reference number: A70500; MUID: 98295987

A: Accession: C70791

A: Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA

A: Residues: 1-810 <COL>

A: Cross-references: GB:AL022121; GB:AL123456; NID: g3261559; PIDN: CAA18004.1; PID: el26454

A: Experimental source: strain H37RV

C: Genetics:

A: Gene: ponA'

Query Match 5.8%; Score 8; DB 2; Length 810;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APPPPAPP 15

|||||||

Db 802 APPPPAPP 809

Search completed: August 16, 2001, 15:45:06

Job time: 296 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2001, 07:35:26 ; Search time 15.24 Seconds
(without alignments)
317.770 Million cell updates/sec

Title: US-09-357-349-4

Perfect score: 139

Sequence: 1 PPQSPRPAPPAPPSPALPR.....VNSTWRTVDRLSATACGCLG 139

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 94743 seqs, 34840360 residues

Word size : 0

Total number of hits satisfying chosen parameters: 94743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 9 | 6.5 | 429 | 1 | GAG-HTLLA |
| 2 | 9 | 6.5 | 429 | 1 | GAG-HTLLC |
| 3 | 9 | 6.5 | 429 | 1 | GAG-HTLLM |
| 4 | 8 | 5.8 | 108 | 1 | NIGM_BOVIN |
| 5 | 8 | 5.8 | 156 | 1 | PSPN_HUMAN |
| 6 | 8 | 5.8 | 430 | 1 | E2F1_MOUSE |
| 7 | 8 | 5.8 | 437 | 1 | E2F1_HUMAN |
| 8 | 8 | 5.8 | 484 | 1 | OAR1_LOCHI |
| 9 | 8 | 5.8 | 484 | 1 | OAR2_LOCHI |
| 10 | 7 | 5.0 | 68 | 1 | BD01_HUMAN |
| 11 | 7 | 5.0 | 68 | 1 | BD01_MACMU |
| 12 | 7 | 5.0 | 161 | 1 | ID4_HUMAN |
| 13 | 7 | 5.0 | 161 | 1 | ID4_MOUSE |
| 14 | 7 | 5.0 | 170 | 1 | VEAR_HCMVA |
| 15 | 7 | 5.0 | 137 | 1 | NRTN_HUMAN |
| 16 | 7 | 5.0 | 207 | 1 | UCR1_CHRVI |
| 17 | 7 | 5.0 | 231 | 1 | RNC_SYNY3 |
| 18 | 7 | 5.0 | 236 | 1 | UL51_PVKKA |
| 19 | 7 | 5.0 | 239 | 1 | WT1_SMIMA |
| 20 | 7 | 5.0 | 242 | 1 | YB11_STRCI |
| 21 | 7 | 5.0 | 260 | 1 | Y757_METJA |
| 22 | 7 | 5.0 | 288 | 1 | WT1_ALLEMI |
| 23 | 7 | 5.0 | 290 | 1 | TRX2_MOUSE |
| 24 | 7 | 5.0 | 299 | 1 | MMS3_MYCTU |
| 25 | 7 | 5.0 | 317 | 1 | TN11_HUMAN |
| 26 | 7 | 5.0 | 319 | 1 | BST1_RAT |
| 27 | 7 | 5.0 | 323 | 1 | PIP-STRCO |
| 28 | 7 | 5.0 | 325 | 1 | PSBO_WHEAT |
| 29 | 7 | 5.0 | 339 | 1 | RLA0_ARCFU |
| 30 | 7 | 5.0 | 361 | 1 | IE63_PVKKA |
| 31 | 7 | 5.0 | 396 | 1 | TCR3_ECOLI |
| 32 | 7 | 5.0 | 431 | 1 | UL61_HCMVA |
| 33 | 7 | 5.0 | 445 | 1 | HISX_MYCSM |

RESULT 1

| ID | GAG-HTLLA | STANDARD; | PRT; | 429 AA. |
|----|---|-----------|------|---------|
| AC | P03345; | | | |
| DT | 21-JUL-1986 (Rel. 01, Created) | | | |
| DT | 21-JUL-1986 (Rel. 01, Last sequence update) | | | |
| DT | 01-OCT-1996 (Rel. 34, Last annotation update) | | | |
| DE | GAG POLYPROTEIN [CONTAINS: MAJOR CORE PROTEINS P19 AND P24; NUCLEIC ACID-BINDING PROTEIN P15]. | | | |
| GN | GAG. | | | |
| OS | Human T-cell leukemia virus type I (strain ATK) (HTLV-I). | | | |
| OC | Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus. | | | |
| OX | NCBI_TaxID=11926; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=83221647; PubMed=6304725; | | | |
| RA | Seiki M., Hattori S., Hirayama Y., Yoshida M.; | | | |
| RT | "Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA."; | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 80:3618-3622(1983). | | | |
| RN | [2] | | | |
| RP | SEQUENCE OF 131-155. | | | |
| RX | MEDLINE=82174582; PubMed=6280175; | | | |
| RA | Oroszlan S., Sargadharan M.G., Copeland T.D., Kalyanaraman V.S., | | | |
| RT | Gilden R.V., Gallo R.C.; | | | |
| RT | "Primary structure analysis of the major internal protein p24 of human type C T-cell leukemia virus."; | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 79:1291-1294(1982). | | | |
| RN | [3] | | | |
| RP | SEQUENCE OF 345-429. | | | |
| RX | MEDLINE=84029174; PubMed=6313426; | | | |
| RA | Copeland T.D., Oroszlan S., Kalyanaraman V.S., Sargadharan M.G., | | | |
| RT | Gallo R.C.; | | | |
| RT | "Complete amino acid sequence of human T-cell leukemia virus structural protein p15."; | | | |
| RL | FEBS Lett. 162:390-395(1983). | | | |
| CC | -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS. | | | |
| CC | -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-POLYPROTEIN. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/ or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; J02029; AAA96672.1; - | | | |
| DR | PIR; A03943; FOLJGH. | | | |
| DR | InterPro; IPR003139; Gag_p19. | | | |
| DR | InterPro; IPR000721; Gag_p24. | | | |
| DR | InterPro; IPR001878; Znf_CCHC. | | | |
| DR | Pfam; PF02228; gag_p19; 1. | | | |
| DR | Pfam; PF00607; gag_p24; 1. | | | |
| DR | Pfam; PF00098; zf-CCHC; 2. | | | |

P47294 mycoplasma
P49552 rattus norv
P19544 homo sapien
P22561 mus musculu
O62651 sus scrofa
P30083 rattus norv
O75949 homo sapien
P47974 homo sapien
Q90643 gallus gall
Q91392 brachydanio
O43516 homo sapien
O63943 mus musculu

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 37-65.
RX TISSUE=Heart;
RC MEDLINE=92389317; PubMed=1518044;
RA Walker J.E., Arizmendi J.M., Dupuis A., Fearley I.M., Finel M.,
RA Medd S.M., Pilkington S.J., Runswick M.J., Skehel J.M.;
RT "Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from
RT bovine heart mitochondria. Application of a novel strategy for
RT sequencing proteins using the polymerase chain reaction.";
RL J. Mol. Biol. 226:1051-1072(1992).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X63216; CAA44901.1; -;
DR PIR; S28241.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transit peptide.
FT TRANSIT 1 36 MITOCHONDRION.
FT CHAIN 37 108 NADH-UBIQUINONE OXIDOREDUCTASE AGGG
FT SUBUNIT.
SQ SEQUENCE 108 AA; 12282 MW; 0F0AD8B6A38120C7 CRC64;

Query Match 5.8%; Score 8; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 ARAAGARG 41
Db ARAAGARG 33

RESULT 5
PSPN_HUMAN
ID PSPN_HUMAN STANDARD; PRT; 156 AA.
AC O60542;
DT 30-MAY-2000 (Rel. 39, Created)
DT 10-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PERSEPHIN PRECURSOR (PSP).
GN PSPN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98150950; PubMed=9491986;
RA Milbrandt J., de Sauvage F.J., Fahrner T.J., Baloh R.H., Leitner M.L.,
RA Tansey M.G., Lampe P.A., Heuckeroth R.O., Kotzbauer P.T.,
RA Simburger K.S., Golden J.P., Davies J.A., Vejsada R., Kato A.C.,
RA Hynes M., Sherman D., Nishimura M., Wang L.-C., Vandlen R., Moffat B.,
RA Klein R.D., Poulsen K., Gray C., Garces A., Henderson C.E.,
RA Phillips H.S., Johnson E.M.;
RT "Persephin, a novel neurotrophic factor related to GDNF and
RT neurturin".
RL Neuron 20:245-253(1998)
CC -1- FUNCTION: EXHIBITS NEUROTROPHIC ACTIVITY ON MESENCEPHALIC
CC DOPAMINERGIC AND MOTOR NEURONS.

CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF040962; AAC39640.1; -;
DR HSP; Q07731; IAGO.
DR MIM; 602921; -;
DR InterPro; IPR001839; TGF-beta.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
KW Growth factor; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 156 PERSEPHIN.
FT DISULFID 66 124 BY SIMILARITY.
FT DISULFID 93 152 BY SIMILARITY.
FT DISULFID 97 134 BY SIMILARITY.
FT DISULFID 123 123 INTERCHAIN (BY SIMILARITY).
SQ SEQUENCE 156 AA; 16600 MW; 6547751653A7044A CRC64;

Query Match 5.8%; Score 8; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 105 PCCRPTRY 112
Db PCCRPTRY 129

RESULT 6
E2F1_MOUSE
ID E2F1_MOUSE STANDARD; PRT; 430 AA.
AC Q61501;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TRANSCRIPTION FACTOR E2F1 (E2F-1).
GN E2F1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS ALBINO;
RX MEDLINE=94158858; PubMed=8114719;
RA Li Y., Slansky J.E., Myers D.J., Drinkwater N.R., Kaelin W.G.,
RA Farnham P.J.;
RT "Cloning, chromosomal location, and characterization of mouse E2F1.";
RT Mol. Cell. Biol. 14:1861-1869(1994).
RN [2]
RP DEVELOPMENTAL EXPRESSION.
RX MEDLINE=98025478; PubMed=9376316;
RA Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L.,
RA Phillips R.A.;
RT "Expression patterns of the E2F family of transcription factors during
RT mouse nervous system development.";
RL Mech. Dev. 66:13-25(1997).
RN [3]
RP FUNCTION IN APOPTOSIS.
RX MEDLINE=98337455; PubMed=9674698;
RA Holmberg C., Helin K., Sehested M., Karlstrom O.;
RT "E2F-1-induced p53-independent apoptosis in transgenic mice.";
RL Oncogene 17:143-155(1998).
CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY

CC WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GGCG,
CC FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS
CC ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE
CC DTF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE
CC PROGRESSION FROM G1 TO S PHASE. E2F-1 BINDS PREFERENTIALLY RB1
CC PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER. IT CAN MEDIATE BOTH
CC CELL PROLIFERATION AND P53-DEPENDENT APOPTOSIS.
CC -1- SUBUNIT: COMPONENT OF THE DTF1/E2F TRANSCRIPTION FACTOR COMPLEX.
CC FORMS HETERODIMERS WITH DP FAMILY MEMBERS. THE E2F-1 COMPLEX BINDS
CC SPECIFICALLY HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1. DURING
CC THE CELL CYCLE, RB1 BECOMES PHOSPHORYLATED IN MID-TO-LATE G1
CC PHASE, DETACHES FROM THE DTF1/E2F COMPLEX, RENDERING E2F
CC TRANSCRIPTIONALLY ACTIVE. VIRAL ONCOPROTEINS, NOTABLY E1A, T-
CC ANTIGEN AND HPV E7, ARE CAPABLE OF SEQUESTERING RB PROTEIN, THUS
CC RELEASING THE ACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DEVELOPMENTAL STAGE: IN THE DEVELOPING NERVOUS SYSTEM, FIRST
CC DETECTED IN THE NEURAL TUBE AT DAY 9.5 DFC. BY DAY 10.5, LEVELS
CC INCREASE THROUGHOUT THE BRAIN, WITH HIGHEST LEVELS IN THE
CC HINDBRAIN AND IN THE SPINAL CORD, EXPRESSED ONLY IN THE POSTRAL
CC HALF. BY DAY 11.5, EXPRESSION FOUND THROUGHOUT THE BRAIN AND
CC SPINAL CORD. FROM DAY 12.5, EXPRESSION RESTRICTED TO THE
CC VENTRICULAR REGIONS OF THE BRAIN, PEAKS AT DAY 13.5 AND DECLINES
CC THEREAFTER. ONLY WEAK EXPRESSION IN THE DEVELOPING SPINAL CORD
CC FROM DAY 11.5-16.5. IN THE DEVELOPING RETINA, EXPRESSION IS
CC CONFINED TO THE UNDIFFERENTIATED RETINOBLASTIC CELL LAYER. IN
CC OTHER DEVELOPING TISSUES, E2F-1 IS EXPRESSED IN KIDNEY, LUNG,
CC LIVER HEPATOCYTES, HEART AND THYMUS. HIGHEST LEVELS IN LIVER.
CC ABSENT IN CHOROID PLEXUS.
CC -1- PTM: PHOSPHORYLATED BY CDK2 AND CYCLIN A-CDK2 IN THE S-PHASE (BY
CC SIMILARITY). BELONGS TO THE E2F/DP FAMILY.
CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L21973; AAA83217.1; -
CC MGD; MGI:101941; E2f1.
CC InterPro: IPR003316; E2F_TDP.
CC Pfam; PF02319; E2F_TDP; 1.
CC KW Transcription regulation; Activator; DNA-binding; Nuclear protein;
CC Phosphorylation; Cell cycle; Apoptosis.
CC FT DOMAIN 62 103 CYCLIN A/CDK2 BINDING (POTENTIAL).
CC FT DOMAIN 105 189 POTENTIAL.
CC FT DOMAIN 148 169 LEUCINE ZIPPER.
CC FT DOMAIN 153 189 DEF BOX.
CC FT DOMAIN 190 279 DIMERIZATION (POTENTIAL).
CC FT DOMAIN 361 430 TRANSACTIVATION (POTENTIAL).
CC FT DOMAIN 402 419 RETINOBLASTOMA PROTEIN RB1 BINDING
CC (POTENTIAL).
CC SEQUENCE 430 AA; 46323 MW; C5DF18AD3B4DDEFA CRC64;

Query Match 5.8%; Score 8; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LLGAGALR 94
|||||||
DB 17 LLGAGALR 24

RESULT 7
E2F1_HUMAN
ID E2F1_HUMAN STANDARD; PRT; 437 AA.
AC Q01094; Q92768; Q13143;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TRANSCRIPTION FACTOR E2F1 (E2F-1) (RETINOBLASTOMA BINDING PROTEIN 3)
DE (RBBP-3) (PRB-BINDING PROTEIN E2F-1) (PBR3) (RETINOBLASTOMA-ASSOCIATED
DE PROTEIN 1) (RBAP-1).
DE GN E2F1 OR RBBP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92346720; PubMed=1638634;
RA Helin K., Lees J.A., Vidal M., Dyson N.J., Harlow E., Fattaey A.;
RT "A cDNA encoding a PRB-binding protein with properties of the
RT transcription factor E2F.";
RL Cell 70:337-350(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92346721; PubMed=1638635;
RA Kaelin W.G. Jr., Krek W., Sellers W.R., Decaprio J.A., Ajchenbaum F.,
RA Fuchs C.S., Chittenden T., Li Y., Farnham P.J., Blarar M.A.,
RA Livingston D.M., Flemington E.K.;
RT "Expression cloning of a cDNA encoding a retinoblastoma-binding
RT protein with E2F-like properties.";
RL Cell 70:351-364(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93078763; PubMed=1448092;
RA Shan B., Zhu X., Chen P.L., Durfee T., Yang Y., Sharp D., Lee W.H.;
RT "Molecular cloning of cellular genes encoding
RT retinoblastoma-associated proteins: identification of a gene with
RT properties of the transcription factor E2F.";
RL Mol. Cell. Biol. 12:5620-5631(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97082961; PubMed=8964493;
RA Neuman E., Sellers W.R.S., McNeil J.A., Lawrence J.B.,
RA Kaelin W.G. Jr.;
RT "Structure and partial genomic sequence of the human E2F1 gene.";
RL Gene 173:163-169(1996).
RN [5]
RP SEQUENCE OF 1-111 FROM N.A.
RX MEDLINE=95047311; PubMed=7958836;
RA Johnson D.G., Ohtani K., Nevins J.R.;
RT "Autoregulatory control of E2F1 expression in response to positive
RT and negative regulators of cell cycle progression.";
RL Genes Dev. 8:1514-1525(1994).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=95140412; PubMed=7838523;
RA Kitagawa M., Higashi H., Suzuki-Takahashi I., Segawa K., Hanks S.K.,
RA Taya Y., Nishimura S., Okuyama A.;
RT "Phosphorylation of E2F-1 by cyclin A-cdk2.";
RL Oncogene 10:229-236(1995).
RN [7]
RP CYCLIN A/CDK2 BINDING DOMAIN.
RX MEDLINE=94306513; PubMed=8033208;
RA Krek W., Ewen M.E., Shirodkar S., Arany Z., Kaelin W.G. Jr.,
RA Livingston D.M.;
RT "Negative regulation of the growth-promoting transcription factor
RT E2F-1 by a stably bound cyclin A-dependent protein kinase.";
RL Cell 78:161-172(1994).
RN [8]
RP DIFFERENTIAL REGULATION BY CYCLIN/CDK2 KINASES.
RX MEDLINE=95047332; PubMed=7958856;
RA Dynlacht B.D., Flores O., Lees J.A., Harlow E.;
RT "Differential regulation of E2F transactivation by cyclin/cdk2
RT complexes.";
RL Genes Dev. 8:1772-1786(1994).
RN [9]
RP REGULATION BY CYCLIN-DEPENDENT KINASES.
RX MEDLINE=97342624; PubMed=9199321;
RA Dynlacht B.D., Moberg K., Lees J.A., Harlow E., Zhu L.;

RT

Specific regulation of E2F family members by cyclin-dependent kinases.*;

RT

Mol. Cell. Biol. 17:3867-3875(1997).

RT

[10]

RP

INHIBITION OF DNA-BINDING.

RT

MEDLINE-95059071; PubMed-7969176;

RT

Xu M., Sheppard K.-A., Peng C.-Y., Yee A.S., Plwnica-Worms H.;

RT

"Cyclin A/CDK2 binds directly to E2F-1 and inhibits the DNA-binding activity of E2F-1/DP-1 by phosphorylation.*";

RT

Mol. Cell. Biol. 14:8420-8431(1994).

RT

[11]

RP

FUNCTION IN APOPTOSIS.

RT

MEDLINE-94224788; PubMed-8170954;

RT

Wu X., Levine A.J.;

RT

"P53 and E2F-1 cooperate to mediate apoptosis.*";

RT

Proc. Natl. Acad. Sci. U.S.A. 91:3602-3606(1994).

RT

[12]

RP

TRANSACTIVATION INHIBITION.

RT

MEDLINE-94019324; PubMed-8413249;

RT

Helin K., Harlow E., Fattaey A.;

RT

"Inhibition of E2F-1 transactivation by direct binding of the retinoblastoma protein.*";

RT

Mol. Cell. Biol. 13:6501-6508(1993).

CC

-1- FUNCTION: TRANSCRIPTION ACTIVATOR THAT BINDS DNA COOPERATIVELY WITH DP PROTEINS THROUGH THE E2 RECOGNITION SITE, TTTC/GCGC, FOUND IN THE PROMOTER REGION OF A NUMBER OF GENES WHOSE PRODUCTS ARE INVOLVED IN CELL CYCLE REGULATION OR IN DNA REPLICATION. THE DRTF1/E2F COMPLEX FUNCTIONS IN THE CONTROL OF CELL-CYCLE PROGRESSION FROM G1 TO S PHASE. E2F-1 BINDS PREFERENTIALLY RB1 PROTEIN, IN A CELL-CYCLE DEPENDENT MANNER. IT CAN MEDIATE BOTH CELL PROLIFERATION AND P53-DEPENDENT APOPTOSIS.

CC

-1- SUBUNIT: COMPONENT OF THE DRTF1/E2F TRANSCRIPTION FACTOR COMPLEX. FORMS HETERODIMERS WITH DP FAMILY MEMBERS. THE E2F-1 COMPLEX BINDS SPECIFICALLY HYPOPHOSPHORYLATED RETINOBLASTOMA PROTEIN RB1. DURING THE CELL CYCLE, RB1 BECOMES PHOSPHORYLATED IN MID-TO-LATE G1 PHASE, DETACHES FROM THE DRTF1/E2F COMPLEX, RENDERING E2F TRANSCRIPTIONALLY ACTIVE. VIRAL ONCOPROTEINS, NOTABLY E1A, T-ANTIGEN AND HPV E7, ARE CAPABLE OF SEQUESTERING RB PROTEIN, THUS RELEASING THE ACTIVE COMPLEX.

CC

-1- SUBCELLULAR LOCATION: NUCLEAR.

CC

-1- PTM: PHOSPHORYLATED BY CDK2 AND CYCLIN A-CDK2 IN THE S-PHASE.

CC

-1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.

CC

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR

EMBL; M96577; AAC35782.1; .

DR

EMBL; S49592; AAB24289.1; ALT_INIT.

DR

EMBL; U47677; AAC50719.1; .

DR

EMBL; U47675; AAC50719.1; JOINED.

DR

EMBL; U47676; AAC50719.1; JOINED.

DR

EMBL; S74230; AAD14150.1; .

DR

PIR; A42998; A42998.

DR

PIR; A42997; A42997.

DR

MIM; 189971; .

DR

InterPro; IPR003316; E2F_TDP.

DR

Pfam; PF02319; E2F_TDP; 1.

KW

Phosphorylation; Cell cycle; Apoptosis.

KW

Transcription regulation; Activator; DNA-binding; Nuclear protein;

FT

DOMAIN 67 108

FT

DNA_BIND 110 194

FT

DOMAIN 153 174

FT

DOMAIN 158 194

FT

DEF_BOX 195 284

FT

DOMAIN 368 437

FT

TRANSACTIVATION (POTENTIAL).

FT

RETINOBLASTOMA PROTEIN RB1 BINDING

FT

(POTENTIAL).

FT

Y->C: NO RETINOBLASTOMA PROTEIN BINDING.

FT

MUTAGEN 411 411

FT

CONFLICT 89 111

FT

CONFLICT 313 313

FT

CONFLICT 322 322

FT

CONFLICT 329 329

FT

CONFLICT 437 437

FT

SEQUENCE 437 AA; 46919 MW; 003B3F654F0C60DF CRC64;

Query Match

Best Local Similarity 100.08; Pred. No. 5.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

87

LLGAGALR 94

Db

19

LLGAGALR 26

RESULT 8

OAR1_LOCOMI

ID

OAR1_LOCOMI

STANDARD;

PRT;

484

AA.

AC

Q25321;

DT

15-JUL-1998 (Rel. 36, Created)

DT

15-JUL-1998 (Rel. 36, Last sequence update)

DT

30-MAY-2000 (Rel. 39, Last annotation update)

DE

TYRAMINE RECEPTOR 1 (TYR-LOC1).

GN

GCRI.

OS

Locusta migratoria (Migratory locust).

OC

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC

Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;

OC

Acridoidea; Acrididae; Acrididae; Locusta.

OC

Acridomorpha; Acrididae; Acrididae; Locusta.

OX

NB1_TaxID=7004;

RN

[1]

SEQUENCE FROM N.A., AND CHARACTERIZATION.

RP

TISSUE-Neuron;

RC

MEDLINE-95279966; PubMed-7760020;

RX

Vanden Broeck J., Vultsteke V., Huybrechts R., de Loof A.;

RA

"Characterization of a cloned locust tyramine receptor cDNA by functional expression in permanently transformed Drosophila S2 cells.*";

RT

J. Neurochem. 64:2387-2395(1995).

CC

-1- FUNCTION: G-PROTEIN COUPLED RECEPTOR FOR TYRAMINE, A KNOWN NEUROTRANSMITTER AND NEUROMODULATOR AND DIRECT PRECURSOR OF OCTOPAMINE. THE RANK ORDER OF POTENCY FOR AGONISTS OF THIS RECEPTOR IS TYRAMINE > NAPHAZOLINE > TOLAZOLINE > DL-OCTOPAMINE > DOPAMINE > EPINEPHRINE > 5-HYDROXYTRYPTAMINE. FOR ANTAGONISTS, THE RANK ORDER IS YOHIMBINE > CHLORPROMAZINE > MANSERIN > PHENTOLAMINE > METOCLOPRAMIDE.

CC

-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC

-1- TISSUE SPECIFICITY: PRESENT MAINLY IN THE CENTRAL NERVOUS SYSTEM, ESPECIALLY IN THE SUPRA- AND SUBESOPHAGEAL, THORACIC AND ABDOMINAL GANGLIA. NOT FOUND IN THE DISTAL PART OF OPTIC LOBES.

CC

-1- DEVELOPMENTAL STAGE: EXPRESSED IN THE NERVOUS SYSTEM BY THE FIRST LARVAL STAGE.

CC

-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR

EMBL; X69520; CAA49268.1; .

DR

GCRCdb; GCR_1763; .

DR

InterPro; IPR000276; GPCR_Rhodpsn.

DR

Pfam; PF00001; 7tm.1; 1.

DR

PRINTS; PR00237; GPCR_Rhodopsin.

DR

PRINTS; PR00564; OCTOPAMINE.

DR

PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.

DR

PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.

KW

G-protein coupled receptor; Transmembrane; Glycoprotein.

FT

DOMAIN 1 54

FT

EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 55 77 1 (POTENTIAL).
 FT DOMAIN 78 87 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 88 109 2 (POTENTIAL).
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 127 147 3 (POTENTIAL).
 FT DOMAIN 148 167 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 168 190 4 (POTENTIAL).
 FT DOMAIN 191 215 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 216 237 5 (POTENTIAL).
 FT DOMAIN 238 411 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 412 433 6 (POTENTIAL).
 FT DOMAIN 434 448 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 449 470 7 (POTENTIAL).
 FT DOMAIN 471 484 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 124 203 BY SIMILARITY.
 FT DOMAIN 286 292 POLY-PRO.
 FT DOMAIN 294 307 ARG/LYS-RICH (BASIC).
 FT DOMAIN 348 360 POLY-THR.
 SQ SEQUENCE 484 AA; 53530 MW; 4484FBC85A810619 CRC64;

Query Match 5.8%; Score 8; DB 1; Length 484;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PPPAPPS 16
 |||||
 Db 286 PPPAPPS 293

RESULT 9
 OAR2_LOCM1

ID OAR2_LOCM1 STANDARD; PRT; 484 AA.
 AC Q25322;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE TYRAMINE RECEPTOR 2 (TYR-LOC2).
 GN GCR2.
 OS Locusta migratoria (Migratory locust).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
 OC Acridomorpha; Acridoidea; Acrididae; Locusta.
 OX NCBI_TaxID=7004;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Neuron;
 RA Vanden Broeck J.;
 RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: G-PROTEIN COUPLED RECEPTOR FOR TYRAMINE, A KNOWN
 CC NEUROTRANSMITTER AND NEUROMODULATOR AND DIRECT PRECURSOR OF
 CC OCTOPAMINE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

EMBL: X59521; CAA49269.1; -
 GCRdb; GCR_1762;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHODOPSN.
 DR PRINTS: PR00664; OCTOPAMINER.
 DR PROSITE: PS00237; G-PROTEIN_RECIP_FL_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECIP_FL_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 54 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 55 77 1 (POTENTIAL).
 FT DOMAIN 78 87 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 88 109 2 (POTENTIAL).
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 127 147 3 (POTENTIAL).
 FT DOMAIN 148 167 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 168 190 4 (POTENTIAL).
 FT DOMAIN 191 215 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 216 237 5 (POTENTIAL).
 FT DOMAIN 238 411 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 412 433 6 (POTENTIAL).
 FT DOMAIN 434 448 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 449 470 7 (POTENTIAL).
 FT DOMAIN 471 484 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 124 203 BY SIMILARITY.
 FT DOMAIN 286 292 POLY-PRO.
 FT DOMAIN 294 307 ARG/LYS-RICH (BASIC).
 FT DOMAIN 348 360 POLY-THR.
 SQ SEQUENCE 484 AA; 53488 MW; 15C10608471D77CD CRC64;

Query Match 5.8%; Score 8; DB 1; Length 484;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PPPAPPS 16
 |||||
 Db 286 PPPAPPS 293

RESULT 10

ID BD01_HUMAN STANDARD; PRT; 68 AA.
 AC Q09753;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE BETA-DEFENSIN 1 PRECURSOR (HBD-1) (DEFENSIN, BETA 1).
 GN DEFB1 OR HBD1 OR BDI.
 OS Homo sapiens (Human), and
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606, 9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA Liu L., Zhao C., Heng H.H.Q., Ganz T.;
 RT "The human beta-defensin-1 and alpha-defensins are encoded by
 RT adjacent genes: two peptide families with differing disulfide
 RT topology share a common ancestry.";
 RL Genomics 43:316-320(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human; TISSUE=Kidney;
 RA Zhao C.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA McCray P.B., Bentley L.;
 RT "Human airway epithelia express a beta-defensin.";
 RN [4]
 RP Am. J. Respir. Cell Mol. Biol. 16:343-349(1997).
 RC SEQUENCE OF 33-68 FROM N.A., AND SEQUENCE OF 33-68.
 RC SPECIES=Human; TISSUE=Kidney, Vagina, and Plasma;
 RX MEDLINE=95354864; PubMed=7628632;

RA Bensch K.W., Raida M., Maegert H.-J., Schulz-Knappe P.,
 RA Forssmann W.-G.;
 RT "hbd-1: a novel beta-defensin from human plasma.";
 RL FEBS Lett. 368:331-335(1995).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-P. troglodytes;
 RA Duits L.A., Langermans J.A.M., van der Straaten T., Vervenne R.A.W.,
 RA Pallansing S., Frost P.A., Hiemstra P.S., Thomas A.W., Nibbering P.H.;
 RT "Expression of beta-defensin 1 in chimpanzee (Pan troglodytes)
 RT airways.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: HAS BACTERICIDAL ACTIVITY (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- TISSUE SPECIFICITY: PLASMA.
 CC -!- MASS SPECTROMETRY: MW=3928; MW_ERR=0.5; METHOD=ELECTROSPRAY;
 CC RANGE=33-68.
 CC -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U50931; AAC51728.1; -;
 DR EMBL; U50930; AAC51728.1; JOINED.
 DR EMBL; X92744; CAA63405.1; -;
 DR EMBL; U73945; AAB49758.1; -;
 DR EMBL; Z50788; CAA90650.1; -;
 DR EMBL; AF188607; AAF04110.1; -;
 DR MIM; 602056; -;
 DR InterPro: IPR001855; Defensin_beta.
 DR Pfam; PF00711; Defensin_beta; 1.
 KW Antibiotic; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 32
 FT CHAIN 33 68 BETA-DEFENSIN 1.
 FT DISULFID 37 66 BY SIMILARITY.
 FT DISULFID 44 59 BY SIMILARITY.
 FT DISULFID 49 67 BY SIMILARITY.
 FT DISULFID 68 7420 MW; BOAB76DEC3B14F94 CRC64;
 SQ SEQUENCE 68 AA; 7420 MW; 5.0%; Score 7; DB 1; Length 68;
 Query Match 5.0%; Score 7; DB 1; Length 68;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 55 GLGHRSD 61
 Db 27 GLGHRSD 33
 RESULT 11
 BD01_MACMU
 ID BD01_MACMU STANDARD; PRT; 68 AA.
 AC 018794;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE BETA-DEFENSIN 1 PRECURSOR (RHBD-1) (DEFENSIN, BETA 1).
 GN DEFB1.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kwok J., Hurllock G., Wu X., Penland C., Wine J.J.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: HAS BACTERICIDAL ACTIVITY (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF014016; AAB66344.1; -;
 DR InterPro: IPR001271; Defensin.
 DR InterPro: IPR001855; Defensin_beta.
 DR Pfam; PF00711; Defensin_beta; 1.
 DR SMART; SM00048; DEFSN; 1.
 KW Antibiotic; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 32 BY SIMILARITY.
 FT CHAIN 33 68 BETA-DEFENSIN 1.
 FT DISULFID 37 66 BY SIMILARITY.
 FT DISULFID 44 59 BY SIMILARITY.
 FT DISULFID 49 67 BY SIMILARITY.
 FT DISULFID 68 AA; 7572 MW; 997336DEC3B0435E CRC64;
 SQ SEQUENCE 68 AA; 7572 MW; 5.0%; Score 7; DB 1; Length 68;
 Query Match 5.0%; Score 7; DB 1; Length 68;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 55 GLGHRSD 61
 Db 27 GLGHRSD 33
 RESULT 12
 ID4_HUMAN
 ID ID4_HUMAN STANDARD; PRT; 161 AA.
 AC P47928; O13005;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DNA-BINDING PROTEIN INHIBITOR ID-4.
 GN ID4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Abdominal adipose tissue;
 RA Kiesling T.L.; Christy B.A.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99087490; Pubmed=9872455;
 RA Rigolet M., Rich T., Gross-Morand M.S., Molina-Gomes D.,
 RA Viegas-Pequignot E., Junien C.;
 RT "cDNA cloning, tissue distribution and chromosomal localization of
 RT the human ID4 gene.";
 RL DNA Res. 5:309-313(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95394461; Pubmed=7665172;
 RA Pagliuca A., Bartoli P.C., Saccone S., della Valle G., Lania L.;
 RT "Molecular cloning of ID4, a novel dominant negative helix-loop-helix
 RT human gene on chromosome 6p21.3-p22.";
 RL Genomics 27:200-203(1995).
 CC -!- FUNCTION: ID (INHIBITOR OF DNA BINDING) HLH PROTEINS LACK A BASIC
 CC DNA-BINDING DOMAIN BUT ARE ABLE TO FORM HETERODIMERS WITH OTHER
 CC HLH PROTEINS, THEREBY INHIBITING DNA BINDING.
 CC -!- SUBUNIT: HETERODIMER WITH OTHER HLH PROTEINS.

CC -|- SUBCELLULAR LOCATION: NUCLEAR.
 CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "ID" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U28368; AAA73923.1; -;
 CC EMBL; Y07958; CAA69255.1; -;
 CC EMBL; U16153; AAA82882.1; -;
 CC MIM; 600581; -;
 CC InterPro; IPR003015; HLH_Myc.
 CC InterPro; IPR001092; HLH_dlm.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
 CC KW Nuclear protein.
 FT DOMAIN 39 48 POLY-ALA.
 FT DOMAIN 65 105 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 118 124 POLY-PRO.
 FT CONFLICT 10 14 SGRKA -> RPLR (IN REF. 3).
 FT CONFLICT 39 40 AA -> Q (IN REF. 3).
 FT CONFLICT 77 79 RLV -> WL (IN REF. 3).
 SQ SEQUENCE 161 AA; 16622 MW; 5814847AE7337339 CRC64;

Query Match 5.0%; Score 7; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 PPPPAPP 15
 Db 118 PPPPAPP 124
 |||||

RESULT 13
 IDA_MOUSE
 AC P41139; STANDARD; PRT; 161 AA.
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DNA-BINDING PROTEIN INHIBITOR ID-4.
 GN ID4 OR ID-4 OR IDB4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP MEDLINE=94188125; PubMed=8139914;
 RA Riechmann V., van Cruichten I., Sablitzky F.;
 RT "The expression pattern of Id4, a novel dominant negative helix-loop-
 RT helix protein, is distinct from Id1, Id2 and Id3."
 RL Nucleic Acids Res. 22:749-755(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99057583; PubMed=9838043;
 RA van Cruichten I., Cinato E., Fox M., King E.R., Newton J.S.,
 RA Riechmann V., Sablitzky F.;
 RT "Structure, chromosomal localisation and expression of the murine
 RT dominant negative helix-loop-helix Id4 gene."
 RL Biochim. Biophys. Acta 1443:55-64(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99051333; PubMed=9831657;
 RA Mancani A., Hernandez M.C., Kuo W.-L., Israel M.A.;

RT "The mouse Id2 and Id4 genes: structural organization and chromosomal
 RT localization."
 RL Gene 222:229-235(1998).
 CC -|- FUNCTION: ID (INHIBITOR OF DNA BINDING) HLH PROTEINS LACK A BASIC
 CC DNA-BINDING DOMAIN BUT ARE ABLE TO FORM HETERODIMERS WITH OTHER
 CC HLH PROTEINS, THEREBY INHIBITING DNA BINDING.
 CC -|- SUBUNIT: HETERODIMER WITH OTHER HLH PROTEINS.
 CC -|- SUBCELLULAR LOCATION: NUCLEAR.
 CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "ID" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X75018; CAA52926.1; -;
 CC EMBL; AJ001572; CAA05120.1; -;
 CC EMBL; AF077859; AAD05213.1; -;
 CC PIR; S43260; S43260.
 CC HSP; P04002; IATF.
 CC TRANSFAC; T01658; -;
 CC MGD; MGI:99414; Idb4.
 CC InterPro; IPR003015; HLH_Myc.
 CC InterPro; IPR001092; HLH_dlm.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
 CC KW Nuclear protein.
 FT DOMAIN 39 48 POLY-ALA.
 FT DOMAIN 65 105 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 118 124 POLY-PRO.
 SQ SEQUENCE 161 AA; 16596 MW; 2DCFF47AF7EE7EED CRC64;

Query Match 5.0%; Score 7; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PPPPAPP 15
 Db 118 PPPPAPP 124
 |||||

RESULT 14
 VEAR_HCMVA
 ID VEAR_HCMVA STANDARD; PRT; 170 AA.
 AC P09694;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE MOST ABUNDANTLY TRANSCRIBED EARLY GENE PROTEIN (TRL4).
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87180235; PubMed=2436392;
 RA Greenaway P.J., Wilkinson G.W.G.;
 RT "Nucleotide sequence of the most abundantly transcribed early gene of
 RT human cytomegalovirus strain AD169."
 RL Virus Res. 7:17-31(1987).
 RN [2]
 RP COMPLETE GENOME.
 RX MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Horsnell T., Hutchison C.A. III, Kourarides T., Martignetti J.A.,
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human

RT cytomagalovirus strain AD169.";
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17403; CAA35453.1; -;
DR EMBL; X17403; CAA35305.1; -;
DR EMBL; M17956; AAA45919.1; -;
DR PIR; S09754; S09754.
KW Early protein.
SQ SEQUENCE 170 AA; 19408 MW; B46C95D4E67DAC35 CRC64;

Query Match 5.0%; Score 7; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PPPAPP 15
Db 24 PPPAPP 30

RESULT 15
NR1N_HUMAN STANDARD; PRT; 197 AA.
AC Q99748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUTURIN PRECURSOR.
GN NEUTURIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97100947; PubMed=8945474;
RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,
RA Crendon D.J., Johnson E.M. Jr., Milbrandt J.;
RT "Neurturin, a relative of glial-cell-line-derived neurotrophic
RT factor.";
RL Nature 384:467-470(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Blum H., Bauersachs S., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP VARIANT HSCR SER-96.
RC TISSUE=Peripheral blood lymphocytes;
RX MEDLINE=98367034; PubMed=9700200;
RA Doray B., Salomon R., Amiel J., Pelet A., Touraine R., Billaud M.,
RA Attie T., Bachy B., Munnich A., Lyonnet S.;
RT "Mutation of the RET ligand, neurturin, supports multigenic
RT inheritance in Hirschsprung disease.";
RL Hum. Mol. Genet. 7:1449-1452(1998).
CC -I- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.
CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT
CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS
CC HAEMOPOIETIC CELLS.
CC -I- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -I- SUBCELLULAR LOCATION: SECRETED.
CC -I- DISEASE: IN ASSOCIATION WITH MUTATIONS OF RET GENE AND POSSIBLY
CC OTHER LOCI, IT IS INVOLVED IN HIRSCHSPRUNG'S DISEASE (HSCR). THIS
CC GENETIC DISORDER OF NEURAL CREST DEVELOPMENT IS CHARACTERIZED BY
CC THE ABSENCE OF INTRAMURAL GANGLION CELLS IN THE HINDGUT; OFTEN

CC RESULTING IN INTESTINAL OBSTRUCTION.
CC -I- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U78110; AAC50898.1; -;
DR EMBL; AL161995; CAB82327.1; -;
DR HSSP; Q07731; IAGQ.
DR MIN; 602018; -;
DR MIN; 142623; -;
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGF-beta.
DR PRINTS; PR00438; GFCYSKNOT.
DR SMART; SM00204; TGF; 1.
DR PROSITE; PS00250; TGF_BETA_1; FALSE_NEG.
KW Growth factor; Signal; Disease mutation; Hirschsprung disease.
FT SIGNAL 1 19
FT PROPEP 20 95
FT CHAIN 96 197
FT CHAIN NEURTURIN.
FT DISULFID 103 165
FT DISULFID 130 194
FT DISULFID 134 196
FT DISULFID 164 164
FT VARIANT 96 96
FT A -> S (IN HSCR; ASSOCIATED TO A RET
FT MUTATION; INCOMPLETE PENETRANCE).
FT /FTid=VAR_009498.
SQ SEQUENCE 197 AA; 22405 MW; 91AFA8C3F8971FD CRC64;

Query Match 5.0%; Score 7; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 QPCCRPT 110
Db 162 QPCCRPT 168

Search completed: August 17, 2001, 07:35:27
Job time: 25 sec

